Journal of Maltese Education Research

Vol:1 No.2 2003 35-59
www.educ.um.edu.mt/jmer

The Physics SEC 2000 Examination: A focus on the Differentiated Paper System

Jacqueline Pace

jacqueline.pace@um.edu.mt

Jacqueline Pace lectures in Physics mainly at the University Junior College ever since its inception in 1995. She is currently involved in the College's set up of its assessment policy and procedures. Before, she taught Physics at the State Sixth Form as well as Physics and Biology at Area Secondary Schools and Junior Lyceums.

Jacqueline Pace graduated in Education in 1991 specialising in Biology and Physics. Her B.Ed. thesis mainly concerned the students' attitudes towards Science. She graduated as Master in Education in 2002 following an individual programme which was centred mainly on assessment and accreditation and also on educational leadership. This paper follows the presentation of her M.Ed. dissertation, The Physics SEC 2000 Examination: A focus on the Differentiated Paper System, which was presented at the Faculty of Education Dissertations Showcase held to mark the 25th Anniversary of the Faculty.

Abstract:

This research centres on Physics SEC examination for May 2000 using MATSEC data. All the different components of this examination were analysed so as to shed light on whether this system is upholding quality, equity and fairness. This study aims at drawing out any distortions that decrease the validity of the system. This research puts forward evidence that an appreciable number of high ability candidates, who should have opted for Paper A, sat for the easier Paper B. This crossover to the softer option distorts the final grade of the candidates. The end result is that, not only the lower ability paper-B candidates are at a disadvantage by being norm-referenced with some very able candidates, but also the high ability candidates undertaking Paper A are finding it harder to access the higher grades since there are fewer candidates of average ability.

1.0 Introduction

The aim of this study is to take a critical look at the differentiated paper system that is adopted locally in the external examinations. However, since the intention was to go into the issue qualitatively besides quantitatively it was decided to narrow this down to one particular subject, and thereby Physics, being the science examination that has the largest candidature was the natural choice. Thus the main concern of this study is to try to bring to the surface the problems, criticisms and implications of the differentiated paper system in Physics SEC examination.

University of Malta, Msida, Malta MSD06
ISSN **1726-9725**

2.0 Aims of the study

The aims of the study stem from the general objective outlined above and are spelled out underneath:

1. How is the fact that more candidates opt for Paper B related to gender or school type? Is there any particular group that is being favoured or hampered by this fact?

2. Does the final grade of the candidates appropriately reflect their ability? Does the paper they had opted for, their gender or the type of school influence their grade? How does this grade distribution relate to the global scores?

3. How does the performance of candidates in each of the constituent papers correlate? Does each constituent part of the examination promote equity of opportunity and outcome for candidates with different gender, paper choice or school type? Or are there discriminations other than that based on ability?

4. What is the impact of the school-assessed coursework, i.e. the Practical, on the final grade and how well does it integrate with the other constituents?

5. In all of the papers, are there any individual items that show gender bias? How does the performance of the paper-A candidates compare with that of the paper-B ones in specific items of the common paper, i.e. Paper 1?

6. In what ways is Paper 2B an easier paper than 2A? Do all of the items of Paper 2 reflect this? And are the cognitive levels under test in line with this? Is Paper 2 well tuned with the ability of the candidates?

7. Can the test items of each constituent paper be considered valid enough on which to base general conclusions regarding the differentiated paper system in Physics SEC?

3.0 Setting

Examinations have an impact on education that goes way beyond the certificate that is assigned. The implications of an invalid examination have a bearing on the development of countless pupils in our schools. Before embarking on an analysis of such a system it is useful to try to map out those areas on which examinations have an impact, especially locally through the SEC system.

4.0 Examinations and their 'Backwash Effect'

The Maltese educational system's emphasis on examinations produces backwash effects on the system itself, on its stakeholders and on society at large. Critics of testing in general, as early as Hoffman (1962) have strongly claimed that examinations are likely to have a number of undesirable effects on a whole network of people. They not only affect the candidates who are about to sit for their examinations; they also influence all the students from the first few days of their

schooling. Black (2001) argues that the examinations are supposed to raise the standards of learning but are not productive in this respect since they do not focus on the actual learning experience but rather on the outcomes of it. This leads to teachers who find themselves coaching their students towards examinations, and this becomes a burden on their pedagogy.

It seems that this pattern is the same all over the world. Min and Xiuwen (2001, p.8) report that due to the examination oriented culture in China,

> "schools and teachers pay closer attention to learning and memorising knowledge instead of training students' skills, attitudes and other non cognitive attributes."

Parents become very much concerned with numbers - the raw scores obtained by their children translate themselves to competitiveness, selection, segregation and streaming. What starts off as an innocent assessment exercise, becomes the arena of social class struggle. One may argue that examinations are being used to perpetuate a society that dominates rather than serves individuals, to channel and fit individuals into categories and separate spaces that lead

> "to a situation where children - and their parents and families - are treated as 'deficit systems' … rather they are victims of **socially created** deficits. And from this we can infer that such deficits can be **socially remedied**."
>
> (Sultana, 1996, pp.120-1)

Examinations have already a commanding influence even on those who opt not to sit for them. And yet not only the examination culture but perhaps more strongly the need to be certified seems to be heavily ingrained in the Maltese society. It is documented that some families used to spend about half a month's income on their child's examination fees (Zammit Mangion, 1992). This was prior to the advent of the local Matriculation and Secondary Examinations Board (MATSEC) when almost all examinations taken were from English GCE Boards. Even though some time has now passed, these fees were even heavier than for the present local ones.

The new SEC system of terminal examinations for certification, however, tried to shift itself towards an assessment approach whereby

> "the introduction of an element of school based assessment in several subjects; relatively low examination fees; avoidance of cultural and gender bias in the examination papers and restriction of registration to candidates who were either in their final year of compulsory schooling or aged 16 or over."
>
> (Ventura and Murphy, 1998, p.47)

In the name of an egalitarian program, one can infer that this new model of assessment focuses more on learning and on holistic development of individuals as learners and has,

> "a constructive focus where the aim is to help rather than sentence the individual; and it emphasises the individual's achievement relative to him or herself rather than to others, or in relation to defined criteria."
>
> (Gipps and Murphy, 1994, p.261)

5.0 Certification by External Examinations

At the end of secondary schooling Form 5 students aged 16 sit for external examinations; either the local SEC examinations provided by MATSEC since 1992 or the foreign GCE examinations, which have still been allowed as an alternative. Although the UK GCSE was not locally introduced, it has been influential to SEC examinations not only in philosophical underpinnings but also in syllabi, structure of operations and system evaluation (Ventura and Murphy, 1998). Thus as with the GCSE, the SEC examination is rather different from the GCE in that it is meant to certify all students finishing secondary schooling rather than solely to give access to post-secondary courses.

Sultana (1998) indicates that while only about 20% of students used to sit for external GCE, "the SEC examination attracts about 80% of the cohort" (p.127). This is evidently an overestimate since the number of candidates varies according to the different SEC examinations. However, what is worth pointing out is that this percentage is large enough as to include students from the lower end of the achievement stream "who previously tended to finish 11 years of schooling without any formal credentials" (*ibid.*).

The SEC examination rates students on a wide range of levels of attainment; 1 being the highest grade and 7 the lowest. Initially, the intention had been to have one graded paper for each subject, with questions of increasing difficulty so as to cater for all abilities. However, the new format SEC examination opted for two 2-hour papers in each subject. Paper 1 is common to all students and thus has to cater for the whole ability range and in most cases includes an aural, oral, practical or coursework component.

Up to the 2001 session, candidates sitting for Papers 1 and 2A qualified for grades 1 to 4 only while candidates sitting for Papers 1 and 2B qualified for grades 4 to 7. Candidates obtaining less than grade 4 from 2A and candidates obtaining less than grade 7 in 2B remain unclassified (U). As from the session of 2002, candidates sitting for Paper 1 and 2A may qualify for a wider range of grades, i.e. from grade 1 to 5 and those below 5 remain unclassified. On the other hand the grading of candidates sitting for Papers 1 and 2B remains unchanged (MATSEC, 2002a).

Registration of all SEC examinations takes place in November, that is, six months before the examination, and no change is allowed after the registration period. The aim of MATSEC to set differentiated papers as Paper 2 in each curriculum subject is as follows:

> PAPER IIA
>
> A paper comprising more demanding questions than those in Paper I - it is designed for the more academically able candidates, and it is targeted at those who expect high achievement and who want to proceed to higher education in the subject
>
> PAPER IIB
>
> A paper comprising less demanding questions than those of Paper I - it is designed for the less academic candidates, and it is targeted at those who do not aspire for high grades
>
> (MATSEC, 2000b)

No change is allowed after the registration period. In May 2000, candidates sitting for Papers 1 and 2A may qualify for grades 1 to 4 only while candidates sitting for Papers 1 and 2B may qualify for grades 4 to 7. Candidates obtaining less than grade 4 from 2A and candidates obtaining less than grade 7 in 2B remain unclassified (U).

In Physics SEC the coursework takes the form of Practical lab reports. An average mark is taken of the best fifteen Practical reports ranging across the Physics curriculum that are worked out by each student throughout the three years of studying Physics. MATSEC markers moderate a wide sample of these to ensure that the marks awarded by the teachers are internally consistent within the school, with the other schools and consistent with the criteria and guidelines that MATSEC puts forward regarding the Practical. The Practical carries 15% weighting of the whole examination.
Paper 1 consists of 10 compulsory fill-in questions to be worked in two hours. Questions range across the Physics curriculum and are allotted 10 marks each summing up to a total of 100 marks. It carries 42.5% of the whole examination and is a common paper for all the Physics candidates.

Paper 2 is also a two-hour paper and carries the remaining proportion of marks, i.e. 42.5%. It consists of 5 extended response questions allotted 20 marks each making a total of 100 marks. This paper tests a narrower content than Paper 1. Candidates either sit for Paper 2A or 2B according to their previous choice when they registered. The 5 questions of both papers concern the same area of Physics; however there are some differences as regards the question format. Paper 2B may have parts of questions that are completely omitted; however it tends to contain more explanatory diagrams than just the text. Paper 2B is also a fill-in paper whereas in Paper 2A the candidates answer on a separate booklet.

6.0 Criticisms of the SEC Differentiated Examination Papers

One of the declared aims of the SEC system, as reported by Ventura and Murphy (1998, p.47) was to avoid "cultural and gender bias in the examination papers". However, there is evidence in local research to suggest that it has only been partially successful in this. Indeed the importation of foreign examinations, lock, stock and barrel as used to happen meant that the cultural and gender biases of those examinations were also perpetrated locally. The British cultural construct of the paper was done away with although the influence of British or American textbooks is still retained especially in the sciences. However, a new local characteristic of cultural bias has been introduced. Indeed it has been shown that the type of school that one attends plays a very important role in the result that one gets from the SEC system (Zammit, 2001).

The gender situation is somewhat less pronounced with mixed reports from different studies. DeBono and Polidano (2001) report that the boys obtained "better grades than the girls" in Mathematics SEC 1999 (p.59). Camilleri (2001) analysed the 2000 session for Mathematics and found "mixed results" with females outperforming males in three of the four papers, i.e. the Core paper, Paper IIA and the Mental (p.117). In an earlier study on Mathematics SEC 1998, Bonanno *et al* (2000) indicated a similar result, since statistical significance was reached in only one paper; boys obtained

better scores in Paper IIB of that session. Zammit (2001) reports that females outperformed males in Chemistry SEC 1999 although not by significant amounts. However, this alternating fate of males and females in the different examinations should not put the gender bias off our agenda. Indeed Zammit also comments that the problem is shifted from the candidates' performance to their subject choice since boys outnumbered girls by far in that Chemistry session. Furthermore, and perhaps more importantly, DeBono and Polidano (2001) found that "boys were found to choose Paper 2A more than the girls" (p.59). The issue of paper choice may also be a variable in the gender scenario.

One of the major queries about the differentiated paper system of students, and may also be of parents, was the practical relevance of grades 6 and 7; are these worth anything or are they equivalent to a fail? A major criticism was offered by the Malta Union of Teachers that affirmed that such a system of differentiated papers requires candidates to declare their choice in advance and this creates undue stress on students, parents and teachers (M.U.T., 2000).

The MATSEC Board took note of this, but it retained the present system of a choice between the two papers since "candidates were likely to get used to making an evaluation of their own abilities before applying for the examination" (Sultana, 1998, p.128). However, Darmanin (1995) argues that this option choice between papers that has to be done along with the registration of the examination traps students, as well as their teachers and parents, in making 'dispositional adjustments' according to beliefs about ability.

> "Choosing options at a time when they are far too young to make informed choices, tending towards gender stereotypes, or as Sultana (1992) has strongly illustrated on a class basis, should make us reconsider why it is we want this early closure." (*Ibid.*, p.86)

MATSEC also argues that the SEC examination is merely implementing the strategy as outlined in the NMC of 1988.

> "As the MATSEC Board brochure advertising the SEC examination declares (p.1), the SEC complements the requirements of the National Minimum Curriculum (1988) by providing a common assessment system of impartial standard, supplying examinations appropriate for students with different abilities, and incorporating recent trends in educational thinking."
>
> (Sultana, 1998, pp.128-9)

MUT countered that still the system of differentiated papers should be abolished and indeed the MUT Conference approved a motion to replace the current system by a common graded paper designed to cover all abilities and certifying between grades 1 to 7 (M.U.T, 2000).

The Examiners' Report for English SEC in 1995 suggests that the choice for two different second papers be abolished and replaced by a properly graded one (MATSEC, 1995). In 1996, the English SEC report refines the suggestion of a "graded paper" by "a single common finely graded paper" (MATSEC, 1996). The Examiners' Report for Physics SEC in 1994 indicated that a significant number of candidates did not choose their paper appropriately (MATSEC, 1994). However, in

1995 the Panel indicated that the candidates seemed to choose more wisely than in the previous year (MATSEC, 1995).

Research about Physics SEC 1996 indicates that some candidates that sat for this examination failed to choose the correct paper (Abdilla et al., 1998). However, these findings still report that there were far more candidates who chose Paper 2A and achieved low scores than there were Paper 2B candidates achieving high scores. This may present a positive perspective to the existence of differentiated papers since they seem to be favouring lower-ability candidates opting for Paper 2B with the possibility of obtaining grades 6 or 7 instead of U. This seems to be in line with the aims of the differentiated papers as explained by MATSEC.

Baldacchino (1998) conducted research into English SEC 1996 and showed that most teachers think, or rather perceive, that students are making the right choices. Of the few who make wrong choices, teachers think this to be due to peer pressure rather than anything else. These teachers also seem to perceive a motivating effect that the choice between papers has on students; mainly that the system did benefit the weaker student.

DeBono and Polidano (2001) performed a study on Mathematics SEC 1999 and argued that "the aims of the MATSEC Board in the introduction of a two-tier paper, are being attained" (p.59). However, this study also reports that some paper-B candidates are also making a wrong choice since their attainment placed them in the top fifth of the whole sample. Another study on Mathematics SEC 2000 done by Camilleri (2001) shows similar evidence. A rather high percentage, about 20%, of paper-B candidates had attained scores that fell in the top 27% of the sample (p.117). Zammit (2001) puts this figure down to about 10% of Chemistry SEC 1999 candidates who "made the wrong paper choice" (p.105).

Research regarding the impact of examinations on Maltese students' achievement and future aspirations shows that a number of teachers and students interviewed were concerned that,

> "in reality the single overlapping grade and the choice of paper was limiting the opportunity of students, labelling them and placing them into categories."
> (Chetcuti and Ventura, 1999, p.6).

One should perhaps ask whether it is not even worse to make the student, not only to accept the label, but choose it himself/herself? Indeed, students attending Trade schools,

> "labelled as low-achieving ever since their primary schooling, have owned and integrated that label to such an extent that they do not believe in their ability to learn and have given up on formal schooling."
> (Sultana, 1996, p.120).

The report prepared by the MATSEC Board for the June 2000 conference on the implementation of the National Minimum Curriculum (Giordmaina, 2000) never hints at how the situation of differentiated papers integrates with the present NMC. This may indicate that the Board believes that the new NMC and the differentiated paper system are not diverging but are indeed complementary. However, one has to say that

MATSEC has just amended the differentiated paper system from the May 2002 session to include two overlapping grades.

7.0 The Sample

The context of the study was chosen to be the May 2000 session of the Physics SEC examination. It was felt that in order to address the research questions, the actual examination scores and scripts needed to be analysed. The data of the performance scores of all the 3939 candidates sitting for this session together with the relevant scripts were reviewed with permission from the MATSEC Board. For ethical reasons, particular data have not been analysed for specific groups that may somehow be identifiable. Candidates who did not complete all the components of the examination have been excluded.

The Junior Lyceums comprise more than 40% of the students who sat for Physics SEC 2000 (Table I). It can also be seen that a larger percentage of Junior Lyceum girls than boys sat for this examination. The next largest proportion of candidates (about 30%) was from the Church schools; however there were more boys than girls.

Area Secondary and Trade schools were combined together since they constituted only 9.1% of the total population of candidates. Although Independent schools made up only 5.3% it was felt that they could not be grouped to other school types since they might produce incongruencies in the overall analysis. Finally, there were the Private and the Post-Secondary candidates who had two common factors; they were not of the same age cohort as the others but were generally older, in some individual cases much older and furthermore they are quite heterogeneous.

Table I: Physics SEC 2000 candidates by Gender and School Category

SCHOOL CATEGORY	BOYS		GIRLS		TOTAL	
	Frq	%	Frq	%	Frq	% tot
Junior Lyceums	657	16.6	1050	26.7	1707	43.3
Area Secondary/ Trade Schools	161	4.1	199	5.0	360	9.1
Church Schools	741	18.8	414	10.5	1155	29.3
Independent Schools	127	3.2	82	2.1	209	5.3
Post-Secondary Schools	97	2.5	160	4.1	257	6.6
Private Candidates	159	4.1	92	2.3	251	6.4
Total	1942	49.3	1997	50.7	3939	100.0

8.0 Methodology

In order to seek answers for the research questions, the Physics SEC 2000 examination was analysed according to three main aspects.

- The distribution of grades and global scores
- The distribution of raw scores in each constituent paper and the correlation between them and with the global score
- Item analysis of each component

Therefore the whole data of the 3939 Physics SEC 2000 candidates was categorised according to gender, age, school type, performance in each examination constituent, in each question (item) as well as according to their final grade and global score. SPSS software was used to carry out relevant statistical analysis; each of the statistical tests used, as mentioned below, were set to test the null hypothesis at 2-tail significance so as to ensure a more rigorous statistic.

9.0 Results

9.1 Cohort Comparisons

More than 75% of the candidates that sat for the examination are born in 1984 (Table II). This means that they had become 16 years of age during the year 2000. The bulk of the rest is presumably composed of candidates who are making a second attempt although this does not exclude that there may be older candidates who are sitting for SEC Physics for the first time. The National Public Registry of 1984 reports 2901 male and 2670 female live births. This means that, of the total males born, only 48.6% sat for Physics SEC 2000 and 59.6% of the females did (an average of 54.1% of the registered population). This is quite a large gap although naturally this statistic does not take into account any fluctuations in numbers due to death, immigration and emigration. Furthermore, some 1984 students may have repeated a school year and so did not sit for the examination despite having the necessary qualifications (16 years old in the year 2000) because they have not completed the entire Physics syllabus.

If one confronts the two main figures that emerge from the Table I, a rather worrying picture emerges. Either only about half of the Maltese population born in 1984 or only around 63% of the entire school-leaving population has actually sat for Physics SEC 2000. Whilst there may still be some that choose another Examination Board and not MATSEC these are not any more in sufficiently large numbers to make an impact. A relatively lower number of candidates sit for the other science subjects, namely Biology and Chemistry (MATSEC, 2002a). Furthermore, there might be a good percentage of these that sit for more than one science SEC examination. Thus the situation may well be that, after 11 years of formal schooling, about a quarter of the school-leaving population remains without any certification in science (also corroborated by data from *ibid*, and MATSEC, 2002c).

Table II: Distribution of Physics SEC 2000 candidates by Gender and Year of Birth

Year of Birth	Boys	Girls	Total
1985	1	2	3
1984	1410	1592	3002
1983	432	311	743
1982	67	75	142
1981	15	11	26
1980 and before	17	6	23
Total	1942	1997	3939

9.2 Paper Choice

Only about one third of the candidates sitting for Physics SEC 2000 took up Paper A (Table III). There is an overall balance between the genders. However, there are school categories where a larger percentage of girls chose Paper A than boys, notably Junior Lyceums and Church schools. But the larger number of boys who chose Paper A in the Independent schools ensures that in general, it is the boys who chose Paper A in a larger percentage.

The traditional view of girls being less confident than boys in science subjects may find some confirmation here (Hili and Zammit [Pace], 1991; Elkjaer, 1987). However, one might also infer that the girls are taking the less risky route, another stereotypical representation of the female gender. One has to see whether this shirking away from the tougher choice does pay off in the end by a higher achievement.

One also notices the rather predictable, very low percentages of candidates who chose Paper A from the Area Secondary and Trade schools. This is especially significant when one recalls that their percentage is only representative of that small part of the cohort (about 20%) that actually sits for the examination (Table II). If the 1.4% (shown in Table III) were to be expressed as a percentage of the cohort of Trade school and Area Secondary school-leaving population it would be a minuscule 0.003%. This is not surprising if one considers that indeed the very existence of Paper B is intimately linked to these candidates. In fact, the differentiated paper system exists because of the tripartite school system. What is perhaps more worrying is that not even half of the Junior Lyceum candidates chose Paper A as opposed to Church and Independent schools whose intake is by and large similar in ability to the Junior Lyceums. This may be indicative of a problem with proper guidance as postulated in the Attitude Survey (Pace, 2000). It may either be that the Junior Lyceum candidates have underestimated their ability or that the Church and Independent schools' candidates have overestimated their own ability. Either way, it is the scores themselves that may shed a light on this.

Table III: Distribution of Physics SEC 2000 candidates' Paper Choice by Gender and School Category

SCHOOL CATEGORY	BOYS		GIRLS		TOTAL	
	Paper A	Paper B	Paper A	Paper B	Paper A	Paper B
Junior Lyceums	143 (21.8%)	514 (78.2%)	316 (30.1%)	734 (69.9%)	459 (26.9%)	1248 (73.1%)
Area Secondary/ Trade Schools	4 (2.5%)	157 (97.5%)	1 (0.5%)	198 (99.5%)	5 (1.4%)	355 (98.6%)
Church Schools	426 (57.5%)	315 (42.5%)	252 (60.9%)	162 (39.1%)	678 (58.7%)	477 (41.3%)
Independent Schools	65 (51.2%)	62 (48.8%)	35 (42.7%)	47 (57.3%)	100 (47.8%)	109 (52.2%)
Post-Secondary Schools	7 (7.2%)	90 (92.8%)	3 (1.9%)	157 (98.1%)	10 (3.9%)	247 (96.1%)
Private Candidates	15 (9.4%)	144 (90.6%)	9 (9.8%)	83 (90.2%)	24 (9.6%)	227 (90.4%)
Total	660 (34.0%)	1282 (66.0%)	616 (30.8%)	1381 (69.2%)	1276 (32.4%)	2663 (67.6%)

9.3 Global Scores and Grades

It is interesting to notice that there are more candidates that manage to obtain grades between 1 and 5 than 6 to U (Table IV). Girls are also more numerous than boys in the upper grades are, and also less numerous in the bottom grades although this difference is not significant overall. However, this gender difference reaches significance in paper-B candidates, in which girls outnumber boys in top grades and vice versa.

Although there are no significant gender differences in the overall distribution of grades, those obtained by the paper-B girls indicate that a good number of them are more able than the respective male counterparts. These girls are obtaining significantly higher grades than boys do (χ^2=8.45, $p<0.01$). This highly contrasts with paper-A candidates who demonstrate a rather opposite effect as the boys outnumber the girls especially in the top two grades.

Table IV. Distribution of Grades into two ability groups by Gender

Frequency N		Paper A	Paper B	Total
Boys	**Grades 1 to 5**	590	611	1201
	Grades 6 to U	70	671	741
Girls	**Grades 1 to 5**	552	736	1288
	Grades 6 to U	64	645	709
X^2 (df=1)		0.02^{NS}	8.45**	2.45^{NS}

NB: ** = p<0.01(2-tailed) ^{NS} = NOT significant

Table V. Distribution of Grades into two ability Groups by School Category

School Category	Ability Groups	
	Grades 1 to 5	**Grades 6 to U**
Junior Lyceums	1122	585
Area Secondary/ Trade Schools	83	277
Church Schools	978	177
Independent Schools	159	50
Post-Secondary Schools	64	193
Private Candidates	83	168
TOTAL	2489	1450

When one analyses the grades that candidates obtained by school, the most prominent result is that the Area Secondary/Trade schools, Private and Post-Secondary candidates have more or less similar results (Table V). On the other hand as regards the Junior Lyceums, Church and Independent schools, these have an opposite skew towards the higher grades. In fact in these three, only about a third of all the candidates fall below the last useful grade, i.e. grade 5.

Thus, the results for the school categories follow expected patterns. One expects that the Junior Lyceums, the Church and the Independent schools obtain the best results, and indeed this is so. On the other hand, the Area Secondary/Trade schools presumably obtain more of the lower grades than higher ones. The Post-Secondary and Private candidates present an element of unpredictability although if one considers that less than 10% of these candidates chose Paper A, then the outcome of grade distribution is not surprising.

9.4 Inter- and Intra- Paper Analysis

In this section an attempt is being made at analysing the various components of the examination both together and separately. From this analysis one deduces some important conclusions. From the inter-paper correlations, it is immediately evident that the Practical does not pull the same weight as the others (Table VI). On the other hand, there is a degree of harmony between all the other papers. However, the apparent similarity may be hiding some important differences between the performance of the different subgroups in each paper.

Table VI. Inter-Paper Pearson r Correlation values by Paper

	Paper 1		
Paper 2A	+0.78	**Paper 2A**	
Practical	+0.32	+0.29	**Practical**
Global	+0.93	+0.94	+0.46
Paper 2B	+0.90	**Paper 2B**	
Practical	+0.33	+0.32	**Practical**
Global	+0.96	+0.96	+0.51

NB: All *r*-values reach significance at $p<0.001$(2-tailed)

When the focus is turned on the Practical, some important findings are made. Firstly, the low correlations albeit rather low, are higher for this session than when they were measured by Xuereb (1996) in a study on Physics Practical work. Furthermore, in that study it was postulated that as regards the Physics Practical, girls were neater and more organised in general but that boys were better at making scientific deductions and at problem solving. These characteristics helped the girls to attain better scores than boys in the Practical, which is also true for this study. This gender difference in the Practical ensures that paper-A girls make up for their tardiness from boys in Paper 1 (Table VII). In general paper-A candidates obtained better Practical scores than paper-B, although this is expected. Moreover, obtaining high scores for their Practical coursework is relatively easy for all candidates and may even be one of the reasons why some candidates opt for Paper A (Figure 1).

Table VII: Mean Scores, s.d. and t-values for each Paper by Gender

	Boys			Girls			Independent samples t-test value
	N	Mean	s.d.	N	Mean	s.d.	
Paper 1	1942	55.17	21.7	1997	54.42	20.7	$+1.10^{NS}$
Paper 2A	660	58.33	14.9	616	58.60	14.1	-0.33^{NS}
Paper 2B	1282	41.82	20.1	1381	44.37	19.4	-3.30***
Practical	1942	12.49	1.9	1997	13.02	1.6	-9.24***

NB: *** = $p < 0.001$ (2-tailed) ^{NS} = NOT significant

Table VIII: Mean Scores and s.d for each Paper by School Category

SCHOOL CATEGORY	Mean Scores (s.d.)			
	Paper 1	Paper 2A	Paper 2B	Practical
Junior Lyceums	54.42 (19.7)	56.97 (14.3)	45.75 (19.4)	13.32 (1.2)
Area Secondary/ Trade Schools	32.52 (18.9)	48.50 (26.8)	30.07 (19.0)	11.08 (2.2)
Church Schools	67.80 (16.5)	60.18 (13.9)	52.14 (17.6)	12.84 (1.6)
Independent Schools	61.12 (17.6)	59.11 (14.5)	50.72 (20.5)	12.73 (1.7)
Post-Secondary Schools	38.57 (12.8)	30.10 (11.0)	33.31 (13.5)	12.06 (2.0)
Private Candidates	40.53 (16.5)	49.94 (18.0)	37.02 (17.1)	11.22 (3.0)

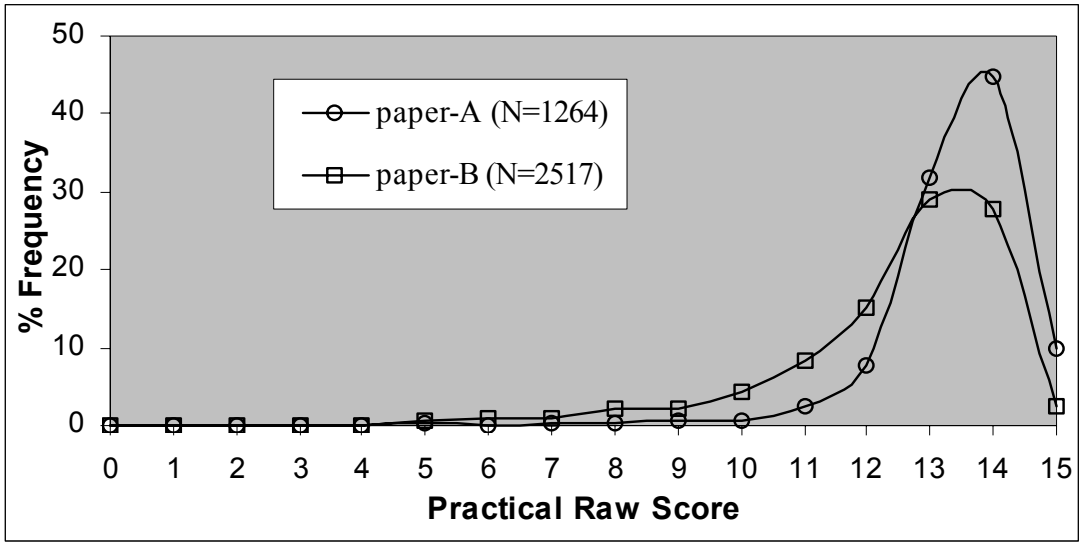
It was also noticed that as regards the Junior Lyceums, their Practical score was better by comparison than their overall performance (Table VIII). This may be indicative that the Practical scores are inflated, although not just for the Junior Lyceums, and they are truly higher on average than those that are corrected by the MATSEC Physics Board, mainly for Private Candidates. It may be that in schools, MATSEC marking schemes may not be properly followed and there have been reports from MATSEC moderation of Physics Practical that there are instances of impression and subjective marking by schoolteachers evident in Practical workbooks.

The Private Candidates are the ones that have their Practical corrected by the MATSEC Physics Board and it is they whom it was found the Practical harms the most. This is not so much due to the Board being more consistent thus apparently stringent in marking than the schoolteachers are but rather because there is a sizeable proportion of them who do not even present their workbook for inspection, thereby

forgoing 15% of the global score. This may be due to these candidates finding difficulties in accessing laboratory resources that are essential for the variation requested in experimental work hence for the Practical workbook to be completed.

As things stand, one might argue that it is rather positive to set part of the examination that is more formative in nature than the rest of the examination papers that inevitably are more summative. Research in Britain indicated that the "strengthening of formative assessment could raise standards of pupil performance" although national surveys reported that there was "very little formative assessment… in science work" (Black, 2001, p.75). It is claimed that Practical coursework presented for Physics SEC is indeed formative assessment. This may be so in the school laboratory but its effect as regards the examination is purely summative as the average mark of fifteen Practical sessions carried out throughout three years of study is taken.

Figure 1. Distribution of Practical Raw Scores by Paper Choice

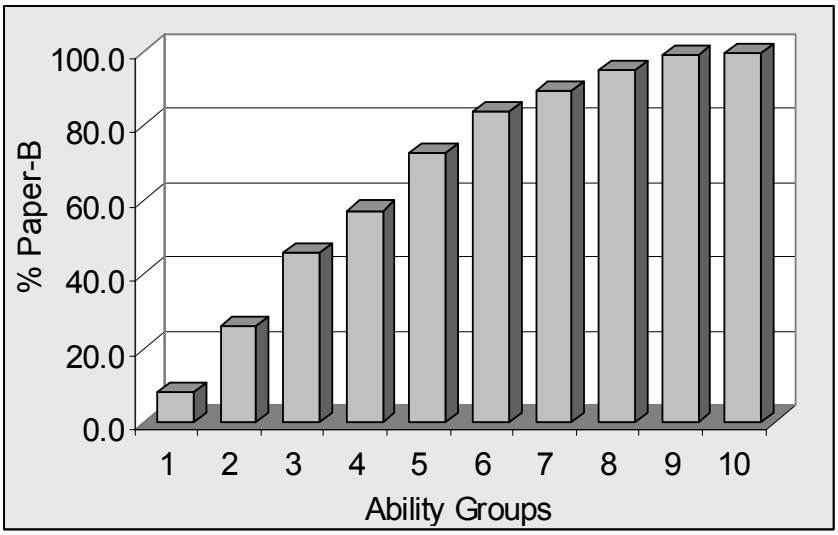


The larger the variation of continuous measures that ultimately sum up the total, the more valid would the final assessment be. However, the Practical seems not to be in tune with the rest of the examination. It seems not to fit within a continuous scale and can cause distortions in the candidates' final grade. Considering that, in general, grade boundaries are basically within the range of 15%; this same percentage assigned to the Practical may thus be influential enough to cause shift of grades.

Reverting now to the other examination papers, when one looks at the performance of paper-A and paper-B candidates in the various papers, there is overwhelming evidence that in Physics SEC, paper-A are superior in ability to paper-B candidates (Table IX). The t-values for all the papers are very large and all significant to the 0.001 level. However, there is ample evidence to demonstrate that there are an appreciable number of high achievers who for some reason or other chose Paper B when they could have aspired at a grade higher than 4 had they chosen Paper A. Even, the standard deviations reveal that paper-A and paper-B candidates oppose each other in homogeneity. The standard deviation might be closer to each other in value if top ability paper-B candidates had chosen Paper A. There is also further confirmation of this from the scores of Paper 1 and Paper 2B.

An appreciable and statistically influential number of paper-B candidates demonstrate very high ability in Paper 1 (Figure 2). This amounts to about a quarter of the candidates that placed in the top three ability groups that had chosen Paper B. Other evidence suggests that there are more girls than boys that compose these paper-B high flyers. However, this does not apply exclusively to girls, as undoubtedly there are similar cases for boys as well.

Figure 2: Percentage paper-B candidates in each ability group for Paper 1



Church and Independent schools seem to have the edge over the other school categories virtually in all the papers. Moreover, given the relatively lower standard deviation registered by these schools in all the different papers (Table VIII), suggests that their candidates are all high achievers. This in turn implies either that all the students in these schools are of a better ability than students in state schools or, more likely, that the low ability students are not even sitting for the examination. This may be due to a number of reasons: either a number of students are being presented for another science subject; or it may also be that they sit for a Physics examination of a foreign Board; or else that the lower ability students in these schools are totally renouncing to attempt this examination, not even opting for Paper B; finally, it may be a mixture of all.

Table IX. Mean Scores by Paper Choice and relevant t-values for each Paper

Mean Scores (s.d.)	Paper-A (N=1276)	Paper-B (N=2663)	Indep. samples t-test
Paper 1 (out of 100)	73.33 (12.3)	45.87 (18.8)	+47.60***
Paper 2 (out of 100)	58.46 (14.5)	43.14 (19.8)	Not Applicable
Practical (out of 15)	13.41 (1.2)	12.43 (1.9)	+16.37***

NB: *** = $p < 0.001$ (2-tailed)

As regards the other categories, the weakness of the Area Secondary/Trade schools is the most prominent feature. It may be argued that the Post-Secondary candidates, having already failed at the examination once, do not have high hopes of obtaining excellent results. As for the Private Candidates, their motivations may be only speculated at since these have little common ground. But the Area Secondary/Trade schools' inability to gain sufficient credibility in any of the papers seems to put the word 'Fail' in bold letters over their candidature. The fact that Paper B has been specifically constructed to meet the needs of these students further amplifies their disappointing performance. On the other hand, had the high ability paper-B candidates, that come mostly from other schools, made the choice that best matches their ability, i.e. Paper A and not the softer Paper B, things could have been very different for them.

Surely their performance in each paper would not have improved but with the more able candidates removed from their scale the performance of the Area Secondary/Trade school candidates would have looked a lot better. Furthermore, the result for paper-A candidates would become presumably more depressed resulting in a higher standard deviation and approximates their distribution to a more normal one. Hence this may improve the result of many paper-A candidates who would either obtain a better grade and more importantly may shift an unclassified one to at least grade 4. All in all, if not always the grades, the actual global scores could have been different for all.

9.5 Item Analysis

One of the important conclusions that can be derived from the item analysis is the validity of Paper 1 as a testing instrument. Indeed there were some problems with the construction of some items in Paper 1, but these problems are not regarded to be serious enough to impinge on the validity of the whole paper (Table X). On reviewing Paper 1 scripts, it was very evident that candidates tend to give only partial answers for descriptive parts of questions where they are allotted 3 marks or more. This is not the case for mathematical or graphical solutions (Pace, 2002). It seems

that the candidates possess the knowledge but are incapable of writing down a full and complete answer. However, the items show an overall balance both in variety of test-items to be answered through a variety of ability skills. Collectively as one whole paper, these test-items are also in equilibrium with the cognitive skills that they test; some testing lower order skills while others insert elements of higher order ones.

Table X: Measures of Central Tendency and Variation for Paper 1 Items

PAPER ONE N=3939	Question (/10)									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Mean	6.06	5.62	6.10	3.90	8.31	5.99	3.82	4.74	4.51	5.76
s.d.	2.8	1.9	3.2	2.5	2.2	3.1	3.1	2.6	2.6	3.3
Skewness	-0.27	-0.29	-0.58	+0.40	-2.52	-0.32	+0.65	-0.04	-0.07	-0.62
Kurtosis	-0.99	-0.08	-0.99	-0.48	+6.12	-1.15	-0.79	-0.92	-1.02	-1.05
***r*-value with Paper 1 total**	0.82	0.66	0.81	0.63	0.55	0.84	0.78	0.78	0.80	0.85

All Pearson's product-moment *r* values are significant at the 0.001 level; 2-tailed.

In general, both Papers 2 have problems with the way they have been constructed. One gets the feeling that Paper 2B is a diluted version of Paper 2A and not a construct on its own that deals with the same topics. Paper 2A does not try to make life easier in any way for the candidate, while Paper 2B is scanty in presenting the problems to its candidates probably out of fear that they might not understand. The end result is that both papers are more difficult than they ought to be. The construction of Paper 2A items should leave more scope for testing higher order skills as in fact it does. Perhaps this should suffice to make Paper 2A more difficult and there would be no reason not to provide fill in spaces. On the other hand, it would help Paper-B candidates if they had the higher order questions presented with better lead-ins including more and better diagrams such that their inherent language difficulties are supplied by visuals that would aid their understanding.

Table XI: Mean scores, s.d. and t-values for Paper 1 Items by Gender

PAPER ONE		Question (/10)									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Boys N=1942	**Mean**	6.13	5.78	6.18	4.08	8.16	6.00	3.78	4.79	4.69	5.61
	s.d.	2.7	2.0	3.2	2.5	2.3	3.2	3.1	2.7	2.7	3.5
Girls N=1997	**Mean**	6.00	5.47	6.01	3.73	8.46	5.97	3.87	4.70	4.34	5.92
	s.d.	2.9	1.8	3.2	2.5	2.1	3.1	3.0	2.6	2.6	3.1
Independent samples t-test value		1.41 NS	5.16 ***	1.68 NS	4.37 ***	-4.27 ***	0.32 NS	-0.93 NS	1.09 NS	4.11 ***	-2.98 **

NB: ** = $p < 0.01$ (2-tailed) *** = $p < 0.001$(2-tailed) NS = NOT significant

Although Paper 1 appeared to be gender neutral as regards the overall score, this does not mean that all the items within this paper carry the same absence of bias (Table XI). Rather it is that the items have gender biases that balance each other out. However, only half of Paper 1 presents a gender bias; three in favour of boys and two against them. Paper 1 thus demonstrates to be a fair test in that it manages to present questions that do relate to the candidates gender-cultural background but without giving any advantages to any one particular gender.

It seems that the same variety of items within Paper 1 has enabled this paper to achieve fairness between genders. From the item analysis, it would also seem that girls prefer to answer questions that are similar in nature to those present in past papers and textbooks. On the other hand, boys have excelled in the questions that called for more creative thought.

However, this trend does not seem to be confirmed in Paper 2 (Table XII). As regards Paper 2A the girls showed that they could be better than boys in highly creative questions. However, on looking closely one notices that this question also calls for the very high answer presentation capabilities. Confirming this trend is the fact that the questions in which boys outperform girls are those that do not require such skills in that these are highly structured questions requiring only short answers or diagrams.

Paper 2B presents a different challenge in that girls outclass boys in four out of five questions (Table XIII). Paper-B girls have better skills than boys in almost all the areas of the syllabus; though the high standard deviations of the boys' distribution of scores for each item indicate that there may well be a good number of high ability boys there as well, apart from the very low ability ones. Finally, all this is highly indicative of the fact that there is a larger proportion of high ability girls more than boys amongst the paper-B candidates. Evidence from the Attitude Survey shows the girls as more likely to take the decision of paper choice upon themselves, undergoing much more stress than boys in choosing between papers (Pace, 2000). This might be indicative of a wrong self-assessment on their part or fear of taking a risk that might backfire.

Table XII: Measures of Central Tendency and Variationfor Paper 2A Items by Gender

PAPER 2A N=1275†		**Question (/20 each)**					
		1	**2**	**3**	**4**	**5**	**Total**
BOYS N=659	**Mean**	13.24	12.32	13.54	8.8	10.64	58.33
	s.d.	3.7	4.3	4.3	3.6	4.5	14.9
GIRLS N=616	**Mean**	12.83	12.66	12.71	10.14	10.26	58.60
	s.d.	3.6	3.7	4.3	3.6	4.1	14.1
Independent Samples t-test values		2.02*	-2.37*	3.42***	-6.70***	1.59^{NS}	-0.33^{NS}
Skewness		-0.37	-0.39	-0.65	+0.16	-0.03	-0.36

NB: † One case is missing
* = $p < 0.05$ (2-tailed) *** = $p < 0.001$ (2-tailed) ^{NS} = NOT significant

Table XIII: Measures of Central Tendency and Variation for Paper 2B Items by Gender

PAPER 2B N=2641†		Question (/20 each)					
		1	2	3	4	5	Total
BOYS N=1273	Mean	10.60	6.87	8.26	6.11	9.98	41.82
	s.d.	4.54	4.42	4.77	4.01	5.62	20.19
GIRLS N=1368	Mean	11.50	7.45	7.85	6.62	10.95	44.37
	s.d.	4.31	4.23	4.57	3.98	5.36	19.41
Independent sample t-test values		-5.26***	-3.44***	+2.30*	-3.25***	-4.54***	-3.30***
Skewness		-0.30	+0.16	+0.16	+0.25	-0.12	-0.04

†22 cases are missing * = $p < 0.05$ (2-tailed) *** = $p < 0.001$(2-tailed)

It is clearly perceptible that there are a number of candidates that are not choosing according to their level of ability. There are undoubtedly a number of candidates that out of fear, or for some reason other than their ability, have chosen to sit for Paper B. The level of skill in the responses of the paper-A candidates to the particularly difficult items in Paper 2A denote that they must be of very high ability (Table XIII). Thus, it is seems that only candidates that have little doubts about their ability in Physics register to sit for Paper A. On the other hand, Paper B seems to be more representative of the whole population in that the candidates there show very diverse abilities.

The overall skewness for Paper 2A is more negative than that for 2B (Tables XII and XIII). This implies that paper-A candidates still find their paper easier than paper-B candidates find theirs. Thus although Paper 2B is indeed an easier paper, this does not mean that it is easier for its candidates to tackle. It is not that paper-B candidates are not capable of answering questions of high levels of cognitive ability. Sometimes parts that require recall are left out and parts that require application and even analysis are answered. It might mean that certain paper-B candidates are not good at studying by rote but still possess a good grasp of Physics. The Examiners' Report 2000 laments that these candidates lack language skills which may also account for their uncertainty in answering recall questions since most of these do require the candidate to state, explain and describe a lot (MATSEC, 2000c).

10.0 Conclusions

The differentiated paper system in Physics SEC is perhaps creating more problems by its existence in the present state than it is solving. It promotes inequitable treatment of the two genders and introduces a possible element of manipulation by the schools that present the candidates. The recent inclusion of grade 5 within the range of grades for Paper A is a move in the right direction (MATSEC, 2002b). The number of

paper-A candidates has already moved up by 10%. However, while this may serve as a temporary solution to the thorny problem of the SEC system, other troubles might be looming ahead. With more formative assessment being advocated by the National Minimum Curriculum (Ministry of Education, 1999), comes a rethinking of the Practical coursework, since as it stands it is a hazard to the validity of the whole examination. One possible solution may be to view the Practical as a source of differentiation and integrate it within Paper 2, whose total percentage is suggested to be 40%. The percentage allotted to the core Paper 1 is thus strengthened to 60% and this is justified by it being fairer and better at assessing candidates' ability than the other Papers. Abolishing the differentiated paper system would seem too big a leap at the moment, although this may be a more equitable option for the future.

References

Abdilla, S., Briffa, N., & Mifsud, D. (1998): *SEC Physics: May 1996***.** Unpublished B.Ed dissertation, Faculty of Education, University of Malta.

Black, P. (2001): Dreams, Strategies and Systems: portraits of assessment past, present and future. *Assessment in Education***,** Vol.8, No.1, pp. 65-85.

Bonanno, V.J., Gatt, K., & Grech, A. (2000): *Analysis of the May 1998 Mathematics SEC Paper.* Unpublished B.Ed dissertation, Faculty of Education, University of Malta.

Camilleri, A. (2001): *Analysis of the May 2000 SEC Examination in Mathematics.* Unpublished B.Ed dissertation, Faculty of Education, University of Malta.

Chetcuti, D. & Ventura, F. (1999): *Maltese Students and Examinations: the impact of differentiated examination papers, gender and type of school on student achievement and future aspirations.* Unpublished paper presented at the 25th Annual IAEA Conf., 23-28 May 1999, Slovenia: Bled.

Darmanin, M. (1995): Closing Options, Opening Opportunities. In Ventura, F. [ed.] *Secondary Education in Malta: Challenges and Opportunities,* Malta: M.U.T. Publication.

DeBono, J. & Polidano, D. (2001): *The May 1999 SEC Examination in Mathematics: A Study.* Unpublished B.Ed dissertation, Faculty of Education, University of Malta.

Elkjaer, J. (1987): Are the girls the problem? In Daniels, J.Z. and Kahle, J.B. (eds.) *Contributions - GASAT 4th International Conference,* Vol. 1, U.S.A: University of Michigan.

Giordmaina, J., [ed.], (2000): *National Curriculum on its Way.* Ministry of Education; Education Division; Faculty of Education, Malta: University of Malta.

Gipps, C. & Murphy, P. (1994): *A Fair Test? Assessment, Achievement and Equity.* Buckingham: Open University Press.

Hili, P. & Zammit [Pace], J (1991): *Attitudes Towards Science: A study of Form 2 pupils.* Unpublished B.Ed dissertation, Faculty of Education, University of Malta.

Hoffman, B.C. (1962): *The Tyranny of Testing.* New York: Crowell-Collier.

Junior College (2000): *Junior College Strategic Plan 2000-2005.* Malta: University of Malta.

Malta Union of Teachers (2000): *Biennial General Conference, Amendments and Motions May-June 2000.* Unpublished report, Malta: M.U.T.

MATSEC (1994): *The Examiners' Reports 1994.* Unpublished report, MATSEC Board, Malta: University of Malta.

MATSEC (1995): *The Examiners' Reports 1995.* Unpublished report, MATSEC Board, Malta: University of Malta.

MATSEC (1996): *The Examiners' Reports 1996.* Unpublished report, MATSEC Board, Malta: University of Malta.

MATSEC (2000a): *The Secondary Education Certificate Regulations 2000-2001,* MATSEC web site. Online. Available: http://home.um.edu.mt/matsec/

MATSEC (2000b): *Physics SEC Examination Papers.* MATSEC Board, Malta: University of Malta.

MATSEC (2000c): *Physics SEC Examiners' Report.* MATSEC Board, Malta: University of Malta.

MATSEC (2000d): *Physics SEC Examination Syllabus.* MATSEC web site. Online. Available: http://home.um.edu.mt/matsec/

MATSEC (2002a): *MATSEC Newsletter,* No. 42, January 2002. MATSEC Board, Malta: University of Malta.

MATSEC (2002b): *SEC Registrations May 2002 compared to May 2001.* MATSEC Board, Malta: University of Malta.

MATSEC (2002c): *SEC Examinations 2001.* MATSEC Board, Malta: University of Malta.

Min, H. & Xiuwen, Y. (2001): Educational assessment in China: lessons from history and future prospects. *Assessment in Education: principles, policy & practice*, Vol. 8, No. 1, March 2001, pp. 5-10.

Ministry of Education (1999): *Creating the Future Together: The National Minimum Curriculum.* Malta: Klabb Kotba Maltin.

National Public Registry (1984): *Data of 1984 Live Births* **(**retrieved in July 2001), Malta: Ministry of Internal Affairs.

Pace, E. & Pace, J. (2002): *School-Leaving Population Sitting for Physics SEC 2000 examination and subsequent Science Related Career Choices.* A paper presented at the CASTME - S & T Education Conference, Malta -10th-13th April 2002.

Pace, J. (2000): *The Differentiated Paper System in the Physics SEC Examination: An Attitude Survey.* Unpublished M.Ed research paper, Faculty of Education, University of Malta.

Pace, J. (2002): *The Physics SEC 2000 Examination: A focus on the Differentiated Paper System.* Unpublished M.Ed dissertation, Faculty of Education, University of Malta.

Sultana, R.G. (1996): Under-Achievement: Are Students failing school or are our schools failing students? In Ventura, F. [ed.] *Secondary Education in Malta: Challenges and Opportunities*. Malta: M.U.T. Publications.

Sultana R.G. (1998): Malta. In Bray, M. and Steward, L. (eds.) *Examination Systems in Small States: Comparative Perspectives on Policies and Operations*. London: Commonwealth Secretariat

Ventura, F. & Murphy, R. (1998): The Impact of measures to promote equity in the Secondary Education Certificate examinations in Malta: an evaluation. In *Mediterranean Journal of Educational Studies*, Vol. 3, No. 1. University of Malta.

Xuereb, P. (1996): *An Evaluation of Physics Practical Work carried out in Secondary Schools.* Unpublished M.Ed dissertation, Faculty of Education, University of Malta.

Zammit, A. (2001): *SEC Chemistry May 1999: A Study.* Unpublished B.Ed dissertation, Faculty of Education, University of Malta.

Zammit Mangion, J. (1992): *Education in Malta.* Malta: Studia Editions.