From Phillips, Brian (Ed.), (1997) Papers on Statistical Education presented at ICME-8 (International Congress on Mathematics Education-8) Seville, Spain, July 14-21, 1996. Hawthorn, Australlia: Swinburne Press. ISBN 0 85590 753 3

BRINGING THE REAL WORLD INTO STATISTICS?

Sharleen Forbes
Statistics New Zealand
New Zealand

Introduction

In the early (primary) years of formal education students usually share a common curriculum. Even at this level though it is unlikely that they share common prior experiences or external environments. Males, females, indigenous peoples and ethnic minorities bring their own view of the world into the schoolroom. This presents a challenge for the teacher who is aware that it is good practice to set learning in familiar contexts. The difficult task is for the teacher to ensure that the learning situation is the child's and not the teacher's.

The division of learning into subject areas is a construct of formal schooling. Specialisation within subjects, and restrictions on the number of subject choices, increases throughout secondary and tertiary education. While mathematics and statistics can be taught and studied purely for their abstract beauty, their power is as tools to solve 'real-world' problems. For statistics in particular, the linking with other subjects (such as biology, social sciences, economics) is critical as techniques appropriate for one area are often inappropriate for another.

However, especially in the senior levels of secondary schooling, students in statistics classes will not have 'other subjects' in common. While this may merely present a challenge for teaching, it is a real concern for equitable assessment.

In the final year of secondary schooling New Zealand students may choose to sit for a national qualification, University Bursary and, until 1990, an additional 'elite' examination, Scholarships. Since 1986, one of the subjects which may be taken is *Mathematics with Statistics* which was originally designed for students needing mathematics in their 'other' subjects, or as a terminating mathematics course.

This report first investigates the subject choices and combinations of male and female Bursary students. A comparison of gender differences in performance in the internally assessed and written examination parts of the 1992 Bursary *Mathematics with Statistics* is provided. Then the statistics questions in all Bursary and Scholarships *Mathematics with Statistics* examination papers from 1986 to 1995 are categorised according to their gender-neutrality, and according to their relationship with other subjects. For those years (1986-1988) for which the data is readily available, the

relationship between gender differences in question selection and performance and question context and content in these papers is also analysed.

Background

University Bursary

In the final year of secondary schooling, the seventh form (year 13) four to six University Bursary papers may be taken. There are two mathematics papers which may be chosen: *Mathematics with Calculus* (primarily designed for students intending to do tertiary mathematics), and *Mathematics with Statistics*. In 1991 (Table 1) *Mathematics with Statistics* was the most popular subject choice for Bursary by males (taken by 63%) and *Mathematics with Calculus* (taken by 48%) was their third choice, with English the second choice. *Mathematics with Statistics* however, was the second choice for females (taken by 49%) after English, and *Mathematics with Calculus* the fifth choice (taken by 31%) following Biology and Geography (Ministry of Education, 1992). This order has changed little over the years with subjects only moving at most one position. For example, in 1995 the order remained the same for males, but for females *Mathematics with Calculus* and Geography were interchanged (Forbes et al, 1996).

Table 1.
The Ten Most Popular Subject Choices In 1991 Bursary Examination

MALES Subject:	Percentage of candidates:	FEMALES Subject:	Percentage of candidates:
Maths with Statistics	63%	English	77%
English	62%	Maths with Statistics	49%
Maths with Calculus	47%	Biology	41%
Physics	38%	Geography	34%
Economics	36%	Maths with Calculus	31%
Geography	33%	History	30%
Chemistry	30%	Economics	28%
Accounting	24%	Chemistry	21%
History	23%	Classical Studies	20%
Classical Studies	21%	Accounting	20%
Total No. Male Candidates	**10627**	**Total No. Female Candidates**	**10983**

New education policies, such as raising the legal minimum leaving age, resulted in large increases during the eighties and early nineties in the numbers of students staying at school until the final year. The retention of females throughout the whole of secondary schooling has shown a dramatic

improvement and now surpasses that of males (by between 300-1200 students annually). In terms of absolute numbers the participation of female students in *Mathematics with Statistics* has steadily increased (from 5045 in 1990 to 5626 in 1995). While numbers of male students fluctuate these have also shown a slight increase (from 6388 in 1990 to 6481 in 1995). Although there is some overlap between the two course prescriptions students may take both mathematics papers. For both males and females the numbers taking both fluctuate and the highest total number of students (6402) occurred in 1992 (Forbes, Blithe, Clark & Chamberlain, 1996).

However, as a percentage of the age cohort at this level of schooling, there has been a dramatic drop in mathematics participation. This drop is more marked for male students, and occurs in both those taking only one mathematics paper, and in those taking both papers. As Table 2 indicates, a smaller percentage of males than females take no mathematics paper. and the proportion in this group is increasing for both genders (Forbes et al,1996). Not surprisingly, the students taking both papers perform better on average than those taking one paper only. This is the case for both genders (Forbes, Blithe, Clark and Robinson, 1990; Morton, Reilly, Robinson and Forbes, 1993). However, a smaller percentage of girls than boys take both papers (Forbes et al, 1996).

Table 2.

Participation in Form 7 (year 13) Mathematics Course

Enrolled in	Percentage of Form 7 (year 13) Females		Males	
	1992	**1994**	**1992**	**1994**
Mathematics with Calculus	34	31	52	46
Mathematics with Statistics	54	44	68	57
Both courses	23	20	40	34
Neither course	35	44	20	31

There are also gender differences in the proportions doing particular maths/science and maths/commerce combinations. In 1986, 36% of male students took at least one maths paper together with either physics or chemistry compared to 18% of female students, and 20% of males took at least one maths paper together with either accountancy or economics compared to 12% of females (Clark, 1989). In 1995 the maths/physics/ chemistry combination had dropped to 21% for males, and 11% for females. The maths/accountancy or economics combinations were taken by 8% of males and 7% of females in 1985. In 1995 over three times as many males as females took *Mathematics with Statistics* together with physics, and one and a half times as many females as males took this paper with biology (Forbes et al, 1996).

Mathematics with Statistics

Mathematics with Statistics is assessed in two parts.

80% external assessment (end-of-year three hour written examination).

20% internal assessment (format left to individual schools).

The internal assessment procedures used vary between schools. Some schools base their internal assessment on the outcome of a single in-depth practical project; some on the amalgamation of a number of small projects. Test results are also included by some schools. In general, however, the internal assessment is based on work in one or more practical projects. The 'project' process is itself a way of bringing the real world into the students' learning of statistics. As students generally can choose their own project topic they have much more freedom to bring their 'real world' into the assessment process rather than be confronted by the 'real world' of the assessor.

Table 3. 1992 Bursary Mathematics with Statistics Results

MARKS	GENDER Male	Female	Total
RAW EXAM MARK			
Mean	54.2	50.7	52.6
Standard Deviation	22.9	21.5	22.3
Number	6215	5059	11274
RAW PROJECT MARK			
Mean	11.8	12.6	12.1
Standard Deviation	4.3	3.7	4.1
Number	6346	5161	11507
COMBINED SCALED MARK (Before inter-subject scaling)			
Mean	53.6	52.3	53.0
Standard Deviation	16.3	14.7	15.6
Number	6346	5161	11507
FINAL MARK RECEIVED BY CANDIDATE (After inter-subject scaling)			
Mean	54.8	53.7	54.3
Standard Deviation	15.9	13.9	15.0
Number	6346	5161	11507

After the internal assessment marks are submitted by each school to the central agency responsible for the awarding of qualifications (New Zealand Qualifications Authority) they are moderated by scaling to have the same mean and standard deviation as that achieved by the school in the written examination. This practice of moderating, or validating, one form of assessment by another has been criticised by some researchers (Cox, 1994; Forbes, 1994). In this case it may mean that the 'real world' project assessment is being moderated by the examination assessment - a narrower and more artificial construct of statistics.

As Table 3 indicates, although the final mark in the *Mathematics with* Statistics paper is a combination of the examination and project marks, it is dominated by the examination performance. The gender difference in means is significant (at the 1% level) for both components, but for the projects marks is in favour of the female students. Thus, on average, female students are out-performing their male counterparts on the internal assessment. The reverse is the case in the written examination, and therefore also for the final marks. This difference may in part result from the different 'other' subject choices of males and females, or perhaps females just prefer the internal assessment. Interestingly, it appears that the gender difference in the mean final mark is reduced by the inter-subject scaling[1].

In order to examine the amount, and type, of 'real world' incorporated into the written examination in this paper it is necessary to examine each question set in some detail.

Analysis of mathematics with statistics questions

Since the subject was introduced in 1986, *Mathematics with Statistics* Bursary and Scholarships examination papers have consisted of two sections of optional long-answer questions. In some years, there was also a set of compulsory short-answer questions. In all papers, one of the long-answer sections contained mathematics questions, the other statistics questions. Students were to select a maximum number of questions from each section.

From 1986-1989, in both Bursary and Scholarships, students could attempt five questions from the two sections but were required to do at least two questions from each section. Previous research (Forbes, 1988) shows that only a small number of students (<5%) did not attempt five questions, so the section chosen for the fifth (final) question can be used as an indicator of preference between the sections. Overall students preferred the statistics section rather than the mathematics section for their final question. However, for all papers, except 1988 Scholarship (when the difference was not significant) a significantly (at 1% level) higher proportion of females than males did so. That is, female students showed a more marked preference for the statistics questions than male students. Since 1989 students have been instructed to choose two questions from each section.

In the period from 1986 to 1989, the gender difference in the mathematics section was also statistically significant and in favour of the male students. The gender difference in the statistics section was always smaller than that in the mathematics section and was not always statistically significant. These results are not surprising considering the high overlap in the mathematics

1 The inter-subject scaling involves first standardising the marks in each subject to the same percentiles. Then each subject is rescaled separately in line with the marks obtained by the students taking that subject, over all their other subjects. It is not clear why this process reduces the gender difference in the *Mathematics with Statistics* paper.

section with *Mathematics with Calculus* and other 'high-numeracy' subjects. One could conjecture that the male students who are doing these subject combinations are being awarded 'more' marks for less new work than other students.

Gender - specific questions

The long-answer statistics questions in all Bursary and Scholarship *Mathematics with Statistics* papers from 1986 to 1995 (inclusive) are categorised in Table 4 according to gender-neutrality and 'other subject' contexts. Questions involving females or female-specific items (for example, pantyhose) or areas (cooking, sewing, etc.) judged to be female specific have been labelled feminine. Questions involving males or male-specific items have been labelled masculine. Neutral questions are sub-classified by whether they are about people, objects (including animals and plants) or are context-free. It should be noted that while an attempt has been made to set questions in a 'real-world' context, these are often minimal and somewhat artificial constructs into which standard questions have been placed. On occasions, however, the context is the driving force behind the whole structure of the question.

As Table 4 indicates there appears to have been a move to have less gender-specific contexts in questions since the Bursary and Scholarship

Table 4. Context of Statistics Questions

	Number of Gender-Specific Questions					
	Gender specific		Neutral			
Paper	Feminine	Masculine	People	Objects/ Games	Context Free	Total
1986 Bursary	1	1	2	1		5
1986 Scholarship	1	1	1	1		4
1987 Bursary			3	2		5
1987 Scholarship	1	1		2		4
1988 Bursary	1		2	1	1	5
1988 Scholarship	1	1	1	1		4
1989 Bursary	1		1 (2x½)	3(2+2x½)		5
1989 Scholarship	1		½	2	½	4
1990 Bursary/School			2	½	½	3
1991 Bursary/School			2⅓	⅓	⅓	3
1992 Bursary/School				3		3
1993 Bursary/School	½		1	1½		3
1994 Bursary/School	1		1	1		3
1995 Bursary/School	1			1½	½	3
TOTAL	9½	4	16⅚	20⅚	2⅚	54

examinations were combined in 1990. The examiner for period from 1986-1989 was female but the gender of the examiners since 1989 is not known. Post-1989, there was a greater tendency for questions to contain parts, with a mix of contexts. Although there is a slight bias in favour of 'female' context questions compared to 'male' context, the large majority of questions are gender-neutral (about objects, games, or people of undefined gender such as students). As Clark (1993) has shown, female students in particular, prefer people-oriented questions.

Relationship with other subjects

Most of the questions were set in contexts which had no apparent relationship to the 'other' subjects which these students could have been studying. Of the total 54 questions, 12 1/3 questions (or parts of questions) were in biology or medical contexts, two in economics, one each in physics or chemistry contexts, three in a pure mathematics (e.g. calculus) context and one was an essay question.

Question preference and performance

Data was available on students' question preferences and performance in individual questions for five papers only; Scholarship in 1986, 1987 and 1988 and Bursary in 1987 and 1988. In these years there was a consistent overall difference in mean marks in favour of males. This difference was statistically significant at the 5% level for both Bursary papers and for the 1986 and 1988 Scholarship paper. In all papers male students had a larger variation (spread of marks) than females. While the mean difference for the mathematics section was consistently statistically significant (at the 5% level) and in favour of males, this was not the case for the statistics section. As stated previously, there was always a smaller gender difference in this section, and the difference was not significant in either the 1986 or 1987 Scholarship papers. In one of these papers the difference was in favour of the female students.

Table 5 gives a breakdown of preference and performance by question context and content in these papers. Overall there are more questions chosen by a higher proportion of female than male students. This is explained, in part, by the greater preference of females to choose a third (final) question from this section rather than from the mathematics section of the paper.

In the Scholarship papers there were significantly (at 5% level) higher proportions of females answering questions 8 in 1986 and 7 in 1987 (both hypothesis testing), and questions 6 in 1987 and 7 in 1988 (both containing sampling). These questions were set in a variety of contexts (students, motorway speeding, pantyhose manufacture and netball). If there is any common denominator it is that they are all about people, supporting Clark's (1993) research referred to earlier. It appears that the content of the question (in this case: standard textbook-type hypothesis testing and sampling which

Table 5(a)

Analysis of Question, Context, Content, Preference and Performance by Gender

– Scholarships papers

Year	Quest.	Context	Content	Proportions Choosing		Mean Mark (out of 20)		Significance (5% level)
				M	F	M	F	
1986				**N = 476**	**N = 158**			
	5	Cycle wheels Male	Normal dist. Conf. int. for mean	51%	50%	13.1	12.6	No
	6	Coin tosses Neutral	Probability	87%	89%	11.2	10.8	No
	7	Dress Manufacture Female	Binomial Approx to binomial	44%	51%	10.3	9.4	No
	8	Rural v Urban Students Neutral	Box plots Hyp. test (means)	66%	79%	9.7	10.7	Yes *
1987				**N = 596**	**N = 228**			
	5	"Craps" Gambling Game Neutral	Probability	54%	52%	8.1	6.7	Yes
	6	Panty Hose Manufacture Female	Expect. Algebra Sampling Probability	52%	62%	7.7	7.5	No
	7	Motorway Speeding Male	Binomial Conf. Int. and Hyp. test (props)	83%	91%	10.5	11.5	Yes *
	8	Larval Tape Worms Neutral (Biology)	Log-fit Hyp. test Poisson	79%	81%	12.5	13.4	Yes *
1988				**N = 640**	**N = 252**			
	5	Lotto (gambling game) Neutral	Probability problem	75%	67%	7.8	5.5	Yes
	6	Context free Neutral (Calculus)	Continuous random vars. Conf. Int. (Mean)	22%	14%	11.6	8.9	Yes
	7	Netball Female	Probability Sampling Hyp. test	69%	79%	8.3	7.8	No
	8	Concrete trucks Male	Normal probs. Hyp. test	89%	91%	14.1	13.9	No

Table 5(b)

Analysis of Question, Context, Content, Preference and Performance by Gender

– Bursary papers

Year	Quest.	Context	Content Choosing	Proportions (out of 15)		Mean Mark		Significance (5% level)
				M	F	M	F	
1987				**N=4796**	**N=3274**			
	7	Production-line Neutral	Probability Expectation	66%	63%	8.2	7.4	Yes
	8	Thistle plants (Biology) Neutral	Poisson	71%	74%	7.9	7.7	Yes
	9	Weekly household income (Economics) Neutral	Sampling Conf. Int. Hyp. test	48%	97%	7.7	7.8	No
	10	Daily alcohol intake (medical) Neutral	Graphs Hyp. test	21%	28%	4.6	5	No
	11	Analgesic medicine (medical) Neutral	Binomial distribution	47%	47%	8.0	7.7	Yes
1988				**N=5755**	**N=4206**			
	7	Birth dates Neutral	Poisson, Binomial Hyp. test	63%	62%	8.7	8.3	Yes
	8	Teething babies Female	Normal distrubution sampling	73%	74%	9.4	8.8	Yes
	9	Biscuit factory Neutral	Probability	32%	32%	6.8	5.5	Yes
	10	English and Maths marks Neutral	Descriptive Statistics	74%	80%	8.2	8.1	No
	11	Context free	Essay on internally assessed project	34%	37%	5.1	5.5	Yes

NOTES:

(1) The Context column indicates a relationship in context with 'other' subject

(2) The significance column indicates that the difference is in favour of females, otherwise it is in favour of males.

generally requires a written paragraph answer) is as least as important as the context. The only question that a significantly higher proportion of males answered was the calculus-based question 6 in 1988. Given that more males than females were concurrently doing *Mathematics with Calculus* this is again not a surprising result.

In the Bursary papers there was a significantly higher proportion of females answering questions 8, 9 and 10 in 1987 and questions 10 and 11 in 1988. Three of these questions were about people, one was in a biology setting and the other was an essay question. Apart from the essay question these were again generally standard questions (on hypothesis testing and descriptive statistics -including graphs). The only question that a significantly higher proportion of males answered was the probability question (7) in 1987 (about a production line). It may be that females were avoiding this question which contained a lot of expectation algebra. Females appear to be choosing questions with "standard" content regardless of the context in which the question was set.

There does appear to be some relationship between question preference and question performance. In all questions where the difference in mean marks was significantly in favour of female students (question 8 in 1986 Scholarship, questions 7 and 8 in 1987 Scholarship, and question 11 in 1987 Bursary), these were chosen by a significantly higher proportion of females than males. The only question for which the female students showed a significantly higher preference than males but there was a significant gender difference in favour of the males was question 8 in the 1987 Bursary paper. In the Scholarship papers, males did significantly better on the probability questions involving gambling games (question 5 in both 1987 and 1988), and the calculus-based question (6) in 1988. In Bursary papers, males did significantly better on questions (7, 8 and 11 in 1987, and 7 and 8 in 1988) involving probability or probability distributions (regardless of the question context).

Summary

When question choice is available in examinations the 'other' subjects taken by students may influence their choices (for example: the female preference for the essay question, the male preference for the calculus question). Although context may also influence question choice, the preliminary investigation above indicates that the content of statistics questions is at least as important as the contexts these are set in. Choice is probably based on a complex interaction between the two. In terms of gender differences, gender-specific preference for a question does appear to be linked with gender-specific performance. Given that there are clear differences between the genders in their subject combinations in the senior secondary school the challenge is to develop interesting and appropriate examination

questions which still provide for equitable assessment of all students at this level.

Internal assessment generally provides some opportunity for students to relate their learning to their 'real world'. As female students perform better when internally assessed than in a written external examination, and vice versa for male students, it would be fairer to use a mix of assessment methods than have the outcome of a statistics course determined by just one of these forms of assessment. Statistics as a discipline has as its basis the quantitative investigation of the real world. We now need to find ways of making the 'real world' of the student, rather than that of the teacher, part of the learning and assessment process.

References

Clark, M. (1989). *Problems in labour force planning.* Seminar, Victoria University (Wider ISOR).

Clark, M. (1993). Friend or foe? *New Zealand Journal of Mathematics, 22* (1), pp 29-42.

Cox, P.(1994). Sex and mathematics CATs (Gender differences in 1993 VCE mathematics assessment). In C. Beesey & D Rasmussen Eds. *Mathematics without limits,* (pp. 20-28). Melbourne: Mathematical Association of Victoria.

Forbes, S.(1988). Mathematics with statistics scholarships examination, *The New Zealand Statistician, 23* (1), pp 17-25.

Forbes, S. (1994). *Assessing statistics learning.* Paper presented to the Fourth International Conference on Statistics Education (ICOTS4), Marrakech, Morocco.

Forbes, S., Blithe, T., Clark, M & Chamberlain, M. (1996). *The continuing need to monitor gender differences.* Paper prepared for the Eighth International Conference on Mathematics Education (ICME8), Seville, Spain.

Forbes, S., Blithe, T., Clark, M. & Robinson, E. (1990). *Mathematics for All? Summary of a study of participation, performance, gender and ethnic differences in mathematics.* Wellington, New Zealand: Ministry of Education.

Morton, M., Reilly, B., Robinson, E. & Forbes, S. (1993). A comparative study of two nation-wide examinations: Mathematics with Calculus and Mathematics with Statistics. *Educational Studies in Mathematics, 00:* pp1-21.

Ministry of Education. (1991). *New Zealand senior school awards and examinations.* Wellington, New Zealand: Demographic and Statistical Analysis Unit, Ministry of Education.