A STUDY OF

THE CANTERBURY - WESTLAND

MATHEMATICS

INTERNAL ASSESSMENT SCHEME

A Research Paper

Presented by Christine Todd

In February 1990

For The Masters of Education Course

At Canterbury University

TABLE OF CONTENTS

ABSTRACT

One factor preventing the wider acceptance of school-based curriculum development and assessment, is the problem of comparing performances of different students, in different schools, in the same and different areas of study.

This research looked at the implications of using a reference test as the moderator, in the Canterbury-Westland Scheme for the internal assessment of Mathematics.

Whilst the study was restricted to one single sex boys school, in a middle class area, the Reference Tests for years 1986 and 1987 were significantly correlated with:

1. the final fifth form rankings for the same year
2. the final sixth form rankings for the following year
3. the Bursary statistics results three years later.

The results of the study supported the moderation procedures used and indicated that the students in the scheme were not advantaged or disadvantaged by being internally assessed, rather than sitting the external examination.

1.0 INTRODUCTION

1.1.0 Background

In 1972 the Minister of Education established a Working Party on Syllabuses in Mathematics at fifth and sixth form level.

The working Party was to consider the mathematical needs of students at this level in relation to their general education and future careers, and to suggest guidelines for the development of syllabuses to meet a wider range of needs than those met by the 1972 examination prescriptions.

In the years prior to the establishment of this Working Party, increasing difficulty had been experienced by teachers of Fifth Form mathematics in assisting their pupils to achieve the aims of the courses, largely because of two factors:

1 The increasing number of pupils remaining at secondary school at least until the end of the third year.

2 The increasing proportion of fifth form pupils who attempted School Certificate Mathematics.

Both of these trends have continued.

School Certificate Mathematics had, for some time, been seen as an unrealistic goal for many Fifth Form students and as early as 1963 a local certificate in mathematics was established in Manukau to cater for low achievers in mathematics.

The introduction of the local certificate courses in mathematics continued to spread, and in 1974 the Canterbury Mathematical Association introduced the Canterbury Mathematics Association (CMA) General Mathematics Certificate. The course was intended for the lower 20% of the "National" Form Five population.

The Working Party concluded that the School Certificate prescription in mathematics had proved satisfactory for about half of the students taking Fifth Form mathematics.

It provided an adequate basis for these pupils to proceed to Sixth Form mathematics.

The working Party stated six basic reasons (from the many possible reasons) why the School Certificate prescription was unsatisfactory for the remaining 50 percent.

1 The range of topics was too great to give slower pupils time for sufficient practice, to allow skills and concepts to become consolidated.

2 Feelings of inadequacy engendered by failure to understand a few of the more difficult topics produced negative attitudes to mathematics, and inhibited the learning of easier content.

3 The pass/fail nature of the examination meant that a large proportion of pupils were aiming at an unattainable goal. This inhibited both learning and teaching and led to "coaching" programmes that were not of lasting benefit to pupils.

4 Limited, or core mathematics courses, for less able pupils proved unsatisfactory, and had been replaced by full mathematics courses, which were not always adapted to the ability or past achievements of these pupils.

5 The proportion of students who took mathematics had increased steadily in recent years. Many of these candidates were studying mathematics because of vocational requirements or parental pressure, or as a prerequisite to further study, rather than from personal choice.

6 Competent mathematics teachers continued to be in short supply in some regions.

The working party found that pupils studying mathematics in Form Five tended to fall into three categories.

1 The group of pupils who entered for local certificates. This probably accounted for only about 20% of all Fifth Form mathematics students.

2 The group of pupils for whom School Certificate Mathematics in its present form was an appropriate goal -

probably less than half of all Fifth Form mathematics students.

3 A large group of students who were regarded as too able to be allowed to enter for local certificates, but for whom the School Certificate Mathematics course proved to be too difficult, and was less relevant than it might be. It was likely that some pupils in this category were at a disadvantage with respect to pupils who could gain good grades in a local certificates, in competition for employment.

At the same time interest in, and a desire for, internal assessment for the award of School Certificate was growing among teachers.

Dissatisfaction with the pass/fail nature of the examination was also expressed.

These views were given additional emphasis, in 1974, by the findings of a special study group on assessment set up by the Educational Development Conference Working Party on Improving Learning and Teaching.

Although a majority of teachers wished the introduction of internal assessment, it was clear that some means of ensuring comparability of standards was also desired.

The working Party on Syllabuses in Mathematics for Forms Five and Six considered that the most effective method of ensuring that Fifth Form mathematics students and their teachers would have realistic and achievable goals to aim for, would be to devise a system of levels for achievement, coupled with internal assessment.

In this way it was hoped that the vast majority of students would be able to learn mathematics which was relevant to them and which was within their grasp, and that teachers would have increased freedom to adjust the syllabus to the needs of their students.

Proposals were therefore put to the School Certificate Examination Board in October 1974 and approval was gained for two experimental schemes to be established in 1975, one in Northland and one in Canterbury-Westland.

Since it was intended that the local groups of teachers should play a significant part in shaping the schemes, the proposals put to them were not detailed. They were, in outline:

1 That the award of School Certificate should be of three levels:

LEVEL 1 Assessment of achievement on the full current School Certificate Mathematics prescription.

LEVEL 2 The current School Certificate prescription modified so that the breadth was limited. Topics omitted, or treated in less depth, were those important only to the study of mathematics as a discipline, and there was to be greater emphasis on mathematics related to practical situations.

LEVEL 3 The level two prescription with breadth further limited, greater emphasis on computation and manipulation, and the application of mathematics to everyday problems.

2 That School Certificate Mathematics be awarded on a system of internal assessment, with a reference (moderating) test to inform schools of how many of their students should receive a particular grade at each of the levels.

A series of meetings were held in Canterbury in February and March 1975, and proposals were put to principals, heads of departments and teachers. It was agreed that a three level scheme for the internal assessment of School Certificate Mathematics should be given a trial.

The scheme got under way in 1975 and 35 schools indicated that they would participate in the scheme and grades and levels were allocated to 4577 candidates.

1975 would have to be regarded as a somewhat atypical year and 1976 would have to be regarded as the first year in which the scheme settled down into what is its present form.

The number of schools involved in the scheme has fluctuated around the range 39-45 schools.

The scheme has been reviewed and evaluated in a number of forms since its inception.

In December 1975, questionnaires were sent to teachers, pupils and parents.

In 1976, studies of the predictive values of the Reference Test were undertaken.

In 1977 Principals of internally assessing schools were invited to comment on the scheme and questionnaires were sent to teachers in the scheme.

Also in 1979 a National In-service Course on all three schemes was held.

1978 saw the first of the School Certificate Examination Board's reviews of the three schemes.

In 1979 a Southern Region In-service Course was held to review the scheme.

The last review was carried out in 1983. Since the 1978 review there had been a number of movements into and out of the scheme. Shirley Boys' High School and Christchurch Girls' High School withdrew before 1978. Waimate High School withdrew in 1981 and Westland High School at the beginning of 1982. The last two schools withdrew for staffing reasons but have continued to retain their contact with the scheme and their pupils still continue to sit the Reference Test.

Schools that have withdrawn and still sit the Reference Test and those who have not participated but sit the Reference Test have been useful in monitoring (in a limited way) the differences that may or may not exist between the distribution of assessment grades between the internal assessment and external examination systems.

The Co-ordinator, Mr Gale, comments that the "general observation from this analysis is both the internal and external assessment schemes produce distributions very close to one another". Variations do occur with a tendency for more students obtaining a "pass" via the internal assessment method but more of the able students gaining more A passes by the external method.

1.2.0 Overall Summary of Scheme as Seen in Schools Visited in 1983

1.2.1 General

Current Operation

The scheme appeared to be functioning smoothly and without fuss in all schools. All schools seemed satisfied with the scheme and its operation. Schools seemed to be under no pressure from pupils, parents or employers due to the internal assessment of School Certificate Mathematics. The scheme has two principal aspects:

1. Internal Assessment.

2. The provision of courses at different levels to suit the ability of the pupils.

All schools seemed to be carrying out the internal assessment aspect with what could be said to be varying levels of success.

However, the observance of, or the attempts to operate the levels aspect varied very widely over the schools visited. In five of the thirteen schools visited a single scheme covered all levels except that some classes seemed to be able to handle more difficult aspects of the set topics. Two schools had some streaming and made some attempts to provide courses at three levels. One school operated at two levels. The other five schools streamed for mathematics and had distinct schemes of work to cover each of the three levels.

It could be said that some schools were not taking full advantage of the opportunities for better mathematics teaching that internal assessment, and the possibilities for the introduction of instruction at different levels with different scheme of work, provided. On the other hand, it seemed that both aspects were being fully utilised to provide the best possible mathematics education for the pupils in Form Five, in some schools, and all the claimed benefits seemed evident.

An interesting feature was the high level of involvement in mathematics in the internally assessed schools. Nearly all of the school seemed to have mathematics as a compulsory subject for all Form Five pupils. although not all were school Certificate candidates.

It was interesting to note that two of the schools had instituted levels systems of teaching for mathematics in Forms Three, Four, and Five, and a third streamed its pupils for mathematics in Forms Three and Four.

A number of the schools noted that involvement in the scheme had produced tremendous professional growth amongst the teachers involved and two principals considered that their schools' involvement in internally assessed mathematics had produced spinoffs into other subject areas and at other levels, particularly form Six.

On the subject of levels and grades a number of points were made. Some schools made it clear that they would welcome the introduction of the "integrated grades' system instead of the present "levels and grades" system for reporting pupil performance.

Another school indicated that it was not happy with the single representative mark for levels and grades, e.g. 90 for level 1A, as it was considered that this had important implications for the new single subject moderation system for

Sixth Form Certificate. Despite claimed benefits, however, two schools claimed that they had problems with pupils. who, as the results of early tests became known tended to give up and presented problems. This complaint was not supported by other schools and so could be a feature of particular schools and a reflection on particular teachers. Other teachers claimed that levels scheme enabled them to keep all pupils working profitably. Again, as has been evident in other recent evaluations, teacher quality is a very important variable. One principal expressed concern about the future operation of the scheme in his school if he were to lose certain key staff members.

Testing programmes varied widely from 50% for End of Year examination, 30% Mid-Year examination, 10% year's work, 10% topic tests to best 5 out of 6 across-the-board topic tests to 60% across-the-board examination plus 40% class teacher input.

The role of the Canterbury Mathematical Association's General Mathematics Certificate (CMA) and its associated prescription is also important.

In some schools the CMA prescription is taken as the level 3 prescription and all CMA candidates (or nearly all) enter for School Certificate Mathematics and obtain a level 3 grade. In other schools some CMA candidates enter for School Certificate or some level 3 School Certificate candidates enter for the CMA examinations. In still other schools either the bottom candidates gain level three grades, or a level three scheme, which is a highly watered-down version of the School Certificate prescription, is used at level 3. CMA Certificate candidates are seen as being below the level of School Certificate and are not entered for School Certificate at all. One school intends to have a Form Five mathematics class in 1984 below the level of the CMA General Mathematics prescription.

1.2.2 PARTICULAR

1. Reference Test

The Reference Test and its placement has always been a significant point of discussion within the Canterbury Mathematics scheme.

The 1983 review revealed the following:

All schools were happy with present timing of the test though one (with experience of Latin Studies and Japanese) said that it could be at the end of the year. One or two teachers mentioned August as a possibility.

The uses made of the test varied. In six of the schools visited the Reference Test result was not used to determine the candidates' final results, i.e. the marks gained were ignored and not "counted-in". In one school the Reference Test could contribute up to 5% of the final mark, in another 10%, in a third 12%, while in the fourth it was counted as one of the year's nine tests (the school dropped the lowest of the 9 marks and used the other 8). Three other schools included a contribution from the Reference Test but did not specify the percentage contribution. Three of the schools used the Reference Test result to adjust the final placement of pupils in classes for the year.

One school, not in the scheme, used the results to determine which of their pupils should be Canterbury Mathematics Association General Mathematics Certificates candidates. (Each year a figure is published and schools are advised that nobody scoring above this figure on the reference test should be a CMA candidate.)

One school used the results of the Reference Test as a basis for interclass comparability.

All schools spent the first month revising for the Reference Test and four indicated that they made positive efforts in Form Four to improve the Reference Test results of their Form Five pupils, particularly those pupils who were the most able, to increase the school's allocation of 1As.

Initially, in the early years of the trial, concentrated revision for the reference test in the hope of getting advantaged results was frowned upon. But with the passage of time there has been a realisation amongst schools that concentrated revision of third and fourth form work at the beginning of the year helps to prepare the pupils for the test (it could be argued for School Certificate Mathematics on March 1) and at the same time the revision of basic work gives a firm foundation for building later work on during the year.

However, there is this new phenomenon which seems to have appeared which involves schools making very intensive efforts in Form Four to improve their Form Five Reference Test results. (SC mathematics in Form Four as described by some). Some schools seemed not to be bothered by other schools indulging in these activities but one of the schools visited was not happy about it.

2 Grade Allocations

"Internal assessment" seems to have varying interpretations. Some would regard internal assessment as implying that the school allocates marks or grades as opposed to some central authority. Others would suggest that internal assessment implies that individual classroom teachers would play, at least, some part (in some cases, a major part) in the allocation of marks or grades to the individual pupils in their classes.

In all of the schools visited internal assessment in the first sense certainly took place, in that the grades allocated to pupils by schools, within the allocations given to them by the coordinator, were those which appeared on their results notice.

However, the methods by which these grades were decided on within the schools varied widely.

In all schools there was evidence of teacher consultation over allocation of grades but at varying levels. In some schools teachers had the right and opportunity to comment once a final ranked list and allocation of grades had been produced. In one school the actual allocation of grades for each class was done by the class teacher on the basis of an allocation of grades handed out by the Head of Department. Cases where these allocations did not fit, classes were adjusted by swapping grades between classes after consultation.

The basis upon which final grades were decided varied widely between schools.

Some of the schools worked on very traditional, mid-year and end-of-year examination systems and the results of these examinations played a very major part in determining the final results of candidates. In some schools the weighting given to these examinations was so high that other contributions could not be expected to play any significant part. However, in one school 40% of the final mark came from individual teacher's class-based tests.

Other schools used a number of across-the-board tests covering either the whole of the Fifth Form or each of the levels, the number of such test varied from six to nine (the lowest mark being dropped).

Three small schools said that they had trouble getting their results to fit into the grade allocations received from the coordinator but both acknowledged that an appeal system existed under which they could seek adjustments to their allocation.

Some schools commented on the wide fluctuations, from year to year, in the grade allocations they received from the co-ordinator. Many felt that there were insufficient level 1A grades to match the ability of their pupils and a number expressed the opinion that many of their pupils classified as 2C (having failed) would have passed the external examination if they had been able to sit.

One small school stated that it had trouble separating its 1As and 1Bs and inserting absentees into the system who missed the Reference Test.

3 Inter-class Comparability

In some schools the question of inter-class comparability is taken care of by having all, or a very high percentage of test or examinations as common (i.e. sat by all).

In some cases schools use a common marking schedule and teachers mark their own pupils. In other schools this was backed up by check marking while in other schools individual teachers mark all the papers in a particular test or examination.

In one school there were three across-the-board tests and six across-each-level tests were moderated against the across-the board tests.

As mentioned previously one school used a 60% common element to establish comparability between classes and allocated a pool of grades to classroom teachers to award on the basis of common and class tests.

A number of schools made no attempt at standardisation or weighting of the results of tests. This could have serious implications for the validity of final rank orderings.

This could have applied to one school which used 6 across-the -board tests and the results of their mid-year examination.

1.2.3 Schools not in the Internal Assessment Scheme

The four schools visited here presented some interesting contrasts. One showed considerable division amongst the staff over internal assessment. The school had been in the scheme for the first two years and had withdrawn. It showed little evidence that it would rejoin the scheme.

In a second school which had initially participated in the scheme but had subsequently withdrawn, the Head of Department agreed with the principles of the internal assessment scheme but no other member of the mathematics staff expressed a willingness to re-enter the scheme until a higher quota of 1A grades could be assured.

In a third school, which had not participated in internal assessment, although its pupils had sat the Reference Test each year, the staff were not strongly opposed to internal assessment but seemed to wish to remain outside the scheme. They liked the high marks their students were gaining in the external examination.

The fourth school had not previously taken part in the scheme but, after long and careful consideration, had come to the decision to take part from 1984 onwards.

All four schools sat the Reference Test in March with the schools taking part in the scheme. The Reference Test was given as one other major reasons why the first two schools did not take part in internal assessment, but the first made use of the results to determine which of its fifth form pupils should be following the CMA general mathematics course instead of sitting the external examination.

In the first school, one criticism of the Reference Test was its position in the school year and it was indicated that a reference test at the end of the Fifth Form year would be more acceptable. The school also expressed the view that, under the present scheme for the moderation of Sixth Form Certificate, the results of the Reference Test in March one year would determine the Sixth Form Certificate grades awarded to and by the school at the end of the following year.

Other facets of the internal assessment scheme objected to by the school included:

- Poor early results tend to cause pupils to give up whereas, with the external examination, there was always hope until the exam has been sat.

- Pupils can get to form 7 without having to sit an external examination in mathematics and are poorly prepared for the Bursaries examination.

- System for indicating results in the internal assessment scheme was too confusing; a need for an integrated grades system was suggested.

- It was felt that as a result of internal assessment the external examination for School Certificate had mellowed to such an extent that the examination was now better suited to the less able.

- Of all school subjects mathematics was the one probably best suited to external examination.

- Internal assessment was too personal; it became a teacher v teacher exercise to pick up the best available grades.

- A reference test at the beginning of the year did not cater for changes of attitude and academic development of students during the fifth form year.

- The school thought its four week intensive revision programme leading up to the external examination was a better way to finish the year.

- The school objected to four weeks of revision at the beginning of the year and felt that is was better to generate pupil interest by starting with something new.

- Teachers had greater flexibility with the external examination as they could, within certain limits imposed by mid-year and end-of-year examinations, teach any topic in any order.

The school indicated that it thought the CMA general mathematics programme a very worth while one. They felt that the papers were ideal for those who would not get much from the School Certificate mathematics examination.

The school indicated that it was not opposed to internal assessment as such (art, Japanese, science, workshop technology are internally assessed at the school) but opposed to the method of internal assessment. They felt that their major objections came from the timing of the reference test and the difference in the number of 1A and 1B equivalents they could gain in the examination. They felt the external examination held pupils' interest better than internal assessment.

The second school presented figures comparing the grades they would have received on the basis of the reference test and the scores gained in the external examination.

	Reference Test	Examination
1979	19 1As	31 As.
1982	14 1As	22 As

These differences were quoted as the major reason why the school did not wish to rejoin the scheme. They claim that their pupils were well prepared and well motivated to sit the Reference Test.

Some of the pupils at this school were involved in the CMA General Mathematics Scheme and the school was very pleased with this scheme.

On the other hand, the school coming into the scheme had discussed the issues amongst the staff, and with Fourth Form parents, and, as a result, the decision had been taken to join the scheme.

The school said that it saw the flexibility of internal assessment as the greatest advantage, although some concern was expressed for the very able students under internal assessment.

The school indicated that initially it would be unlikely to operate in levels but they were keen to get their CMA general mathematics candidates under the School Certificate "umbrella".

The school indicated that they were very well aware of the importance of the reference test.

1.3.0 Comments on the Scheme in 1989

This scheme is now entering its 15th year which must in itself be testimony to it being a valid method in assessing and grading students' mathematical ability at this level. During the past years there has been some minor changes to its operation. The most significant being the EXTERNAL Exam's system of grading, accepting the schemes own gradings of A1, A2, B1, B2 etc.

It is not now just a Canterbury Scheme but a Canterbury-Westland-Southland Scheme, with two Schools in Southland joining (Menzies and Northern Southland) and a Dunedin High School keen to join next year. With all this support perhaps questions should be asked to those schools still pursuing the external exam system, as to what do they see so wrong with Internal Assessment - especially with the demise of School Certificate in its present form around the corner.

The following comments are put together by Jenny Campbell (HOD Maths, Northern Southland College) who spent three days at the end of 1988 visiting some local schools before a decision was made for her school to enter the scheme.

1. All teachers were very enthusiastic about the scheme, and indicated no intentions to change.
2. All indicated they did not gear the end of Form Four year to the Reference Test, but that in February time was important - it was essential to get maximum class contact.

3. Several had made up their own revision booklets or worksheets to use during February while others used the Canterbury Mathematics Association booklet.

4. Schools were happy with the Reference Test (R/T) on the 1st March - later than that would mean students would get bored with revision.

5. All were happy with not being tied to School Certificate prescription for the rest of the year.

6. Most used the reference test in some form for an assessment. Most had common exams. Several also had common tests, others didn't but used the Reference Test to standardize other tests so there was inter-class moderation.

7. Some were very particular about students who were absent for sickness etc. while others let them sit the test later or gave them an assessed mark.

8. Nearly all schools were happy with their grades - even top ones!

9. Only a few give the actual Reference Test papers back to the students but all gave the marks back - confidentially.

10. All stressed that the R/T gains grades for the school - not for the individual - students need to be reminded

and have it explained quite often.

11. Grades on reports are done on a very special basis with no real reference to possible School Certificate grades.

12. Several schools see that tests early in the year are easier and have a high median to give the students a positive feeling about Mathematics and give them a sense of success and achievement.

13. Students never see final marks or grades before School Certificate results.

14. All felt it was fairer and more accurate to have internal assessment rather than have a three hour exam and have to produce all your knowledge at the one time.

15. All schools said teachers could use projects, assignments etc. but it was up to the individual teacher whether they used it for an assessment mark or not - not many did.

16. Most schools had 'essential' chapters in their scheme mentioned, and others as optional.

17. Letters were sent to parents if a pupil missed tests and it appeared suspicious.

18. Meetings to discuss problems included students and parents.

2.0.0 PURPOSE OF THE STUDY

The purpose of the research was to make a pilot study of some of the criticisms expressed by the schools not in the Canterbury-Westland Scheme.

Most of the criticisms involved the Reference Test, including its position in the school year; its use as a moderator for the final Fifth Form grades and hence the following year's Sixth Form Certificate grades; and the distribution of grades allocated to the schools - in particular the shortage of 1A grades.

2.0.1 Allocation of Grades to the Canterbury-Westland School Certificate Mathematics Scheme

Research done by the Department of Education Assessment Division [1] shows that there is a significant relationshop between School Certificate Mathematics grades and the grades obtained in all other School Certificate subjects combined. This correlation forms the basis for the allocation of grades to the Canterbury-Westland Scheme by the Department.

SCHOOL CERTIFICATE MEANS

SUBJECT	1987	1986	1985
English	49.4	49.2	49.3
Econ. Stud.	52.2	52.0	51.4
French	60.7	61.0	61.2
Geography	49.4	49.5	50.5
Home Econ.	44.1	44.6	45.1
Math (045)	52.2	52.5	55.8
Maths (026)	52.2	52.0	51.4

(Data is for all candidates, scaled.)

Table 1

[1] Information obtained from Trevor Boylr of the Department of Education Assessment Division.

Each year the distribution of grades in all other School Certificate subjects are compiled for the preceding four years for both Canterbury-Westland internally assessed schools (MX045) and for "all other areas" (MX026). This provides the comparative profile of the average ability of the Canterbury-Westland students to non - Canterbury-Westland students.

Table 2 shows this distribution for 1983, 1984, 1985 and 1986.
The table also shows the difference in the number of each grade obtained by both schemes for these years. These figures are used to calculate the mean of the differences for the four years.

CANTERBURY-WESTLAND SC MATHEMATICS AND NON-CANTERBURY SC MATHEMATICS

OTHER SUBJECT MARKS - PERCENT IN GRADE

	D 0-15	C2 16-30	C1 31-44	B2 45-55	B1 56-67	A2 68-79	A1 80-100	Grade
1986	0.5	6.6	19.6	25.5	26.7	15.4	5.7	Mx 045 o/s
	0.8	7.6	20.3	25.6	25.8	14.5	5.4	Mx 026 o/s
	-0.3	-1.0	-0.7	-0.1	+0.9	+0.9	+0.3	045-026 o/s
1985	0.4	7.5	19.6	25.9	26.2	15.1	5.3	Mx 045 o/s
	0.9	8.1	20.6	25.3	25.8	14.1	5.2	Mx 026 o/s
	-0.5	-0.6	-1.0	+0.6	+0.4	+1.0	+0.1	045-026 o/s
1984	0.6	6.7	20.1	26.7	25.8	14.9	5.2	Mx 045 o/s
	0.8	7.9	20.8	25.8	25.5	14.1	5.1	Mx 026 o/s
	-0.2	-1.2	-0.7	+0.9	+0.3	+0.8	+0.1	045-026 o/s
1983	0.6	7.8	21.8	26.1	25.0	13.8	4.9	Mx 045 o/s
	0.6	7.4	21.2	26.4	25.0	14.3	5.1	Mx 026 o/s
	0.0	+0.4	+0.6	-0.3	0.0	-0.5	-0.2	045-026 o/s
MEAN	-0.25	-0.60	-0.45	+0.28	+0.40	+0.55	+0.08	045-026 o/s

Table 2

Table 3 shows the distribution of grades obtained by non - Canterbury-Westland Scheme students in 1983, 1984, 1985 and 1986.

026 MATHS - PERCENT IN GRADE

	D	C2	C1	B2	B1	A2	A1
1986	0.8	11.0	23.0	25.5	22.4	12.6	5.0
1985	1.0	10.6	21.2	27.4	22.1	12.8	4.9
1984	1.1	10.7	21.6	27.1	22.0	13.1	4.4
1983	1.1	9.6	22.7	27.1	22.3	12.6	4.6
MEAN	1.00	10.48	22.13	26.70	22.20	12.78	4.73

Table 3

The mean number of grades for the non - Canterbury-Westland scheme students and the mean differences from table 2 are used to determine the distribution of grades for the Canterbury-Westland scheme.

Table 4 shows the distribution of grades and tolerances allocated to the Canterbury-Westland Scheme in 1987.

TARGETS FOR CANTERBURY-WESTLAND SC MATHEMATICS 1987

Percentage given grades and tolerances

	D	C2	C1	B2	B1	A2	A1	
	0.8	9.9	21.7	27.0	22.6	13.3	4.8	
+/-	0.3	1.4	2.0	2.4	0.4	0.6	0.6	+/-

Table 4

The tolerances are based on a 95% confidence interval, 3 degree t-test on the other areas over the last 4 years.

Table 5 shows the actual grades allocated by the scheme in 1987.

D	C2	C1	B2	B1	A2	A1
0.4	9.6	21.3	27.5	22.8	13.7	4.8

Table 5

This is basicially the same analysis used with each School Certificate subject. The difference being that one is based on the current year's results, the other on the cohorts average performance over the previous years.

Tables 6, 7, 8, 9, 10, and 11 show the distribution of grades allocated to the scheme and the school used in this study in 1986 and 1987 - based on the pupils performance in the reference test.

CANTERBURY - WESTLAND SCHOOL CERTIFICATE MATHEMATICS

A. 1986

1. SUMMARY OF LEVELS AND GRADES

	No. of Candidates	Test Total	Mean	Median	St. Dev.
March Return of R/T	6093	50	23.5	23	11.4
August Return of R/T	5114	50	25.6	25	10.8

Table 6

2. ALLOCATION OF LEVELS AND GRADES

	GRADES	A1	A2	B1	B2	C1	C2	D	TOTAL
For all Schools	No.of Candidates	258	678	1174	1357	1132	479	36	5114
	%	5.0	13.3	23.0	26.5	22.1	9.4	0.7	
For the School in this Study	No.of Candidates	37	64	60	42	21	2	0	226
	%	16.4	28.3	26.5	18.6	9.3	0.9		

Table 7

3. SCHOOLS' FINAL DISTRIBUTION OF GRADES

GRADES	A1	A2	B1	B2	C1	C2	D	TOTAL
	38	65	62	46	22	2	0	235

Table 8

B. 1987

1. SUMMARY OF LEVELS AND GRADES

	No. of Candidates	Test Total	Mean	Median	St. Dev.
March Return of R/T	6347	60	27.4	26	12.5
August Return of R/T	5428	60	29.4	29	11.9

Table 9

2. ALLOCATION OF LEVELS AND GRADES

	GRADES	A1	A2	B1	B2	C1	C2	D	TOTAL
For all Schools	No.of Candidates	253	719	1214	1484	1182	547	29	5428
	%	4.7	13.2	22.4	27.3	21.8	10.1	0.5	
For the School in this Study	No.of Candidates	38	69	58	43	38	83	1	255
	%	14.9	27.1	22.7	16.9	14.9	3.1	0.4	

Table 10

3. SCHOOLS FINAL DISTRIBUTION OF GRADES

GRADES	A1	A2	B1	B2	C1	C2	D	TOTAL
	38	70	58	42	35	7	1	251

Table 11

2.0.2 Allocation of Sixth Form Certificate Grades

In 1987 some schools noticed an apparent anomaly in the marks used by the Education Departments Computer to generate Sixth Form Certificate (S.F.C.) grades from the student's School Certificate (S.C.) Mathematics grade obtained the previous year.

There was concern from Schools at the end of that year that the Canterbury Mathematics Scheme did not generate any Grade 2's for S.F.C. (Actually the Scheme did not generate any 4's or 6's either). E.g. A Canterbury student gaining an A2 would have in the past been given 74% which would have earned a

S.F.C. Grade 3. However, an external Mathematics student gaining an A2 in S.C. with a score of 76 would have earned a Grade 2. The Department was also concerned that the Canterbury Scheme generated too many S.F.C. "1's" which distorted the grade pools going to schools. This was because the S.C. mark read was the median mark of that S.C. grade and these did not always line up with the mark divisions used to generate S.F.C. grades.

e.g. S.C. Grade	D	C2		C1		B2		B1		A2		A1		
lowest mark of boundary		16		31		45		56		68		80		
median score			24		38		51		62		74			85
lowest boundary mark for 1988 SFC grade					34	45		54	61		69	76	83	
SF grade			8		7	6		5	4		3	2	1	

But it was not as simple as this because all students scoring say 83% in S.C. would earn their school a level 1 Grade for S.F.C. A process of "sequence-proportioning" between grades was thus carried out throughout the whole range of marks and the degree of adjustments made at each level varied from year to year in grade to grade. This adjustment did not only affect the Internally Assessed S.C. Mathematics scores but all results in all subjects.

This situation was unsatisfactory given the growing prominence the S.F.C. Award was gaining.

After discussions with the Qualification and Assessment unit there were two alternatives.

1. For schools to assign marks to the students as well as their grade. This was discarded as it was working against The Principle of the Scheme i.e. to grade students in levels - or as close as possible.

2. For the pool of A1 grades to be split in a ratio that would generate as close as possible the required distribution of S.F.C. Grades 1, 2, 3, etc. as determined by the Mathematics students performance in their 'other' subjects. This was found to be the preferred option.

The following process is thus applied in each school. The computer assigns the mark of 87 to the first student (alphabetically) to obtain an A1 in S.C. Mathematics (Canterbury). The second student a 78, the third student 87. Then this cycle is repeated. An analysis of this showed that approximately two thirds of the A1 students in Mathematicss will earn for their school a S.F.C. "1" and one third will earn for their school a "2". The ratio chosen for splitting the grades between the higher and lower grades is derived from the S.C. results of the previous year. In cases of doubt the benefit is given to the school.

The first purpose of the research reported here is to shed additional light on the controversy created by the distribution of grades by investigating whether there is any correlation between the reference test rankings and the final fifth form rankings and the final sixth form rankings the following year. In addition the relationship between the reference test marks and the Bursary Statistics marks two years later was also investigated.

Very few teachers in the internally assessed schools are now in a position to compare internal assessment for School Certificate mathematics with the external examination because, since the scheme is in its fifteenth year, few teachers have had the experience of preparing pupils for the external examination. Nevertheless in the 1983 review some teachers in the scheme as well as the teachers not in the scheme thought that they could get more borderline pupils through if they were being externally examined and that they could earn a higher percentage of "1A" passes. However because I taught mathematics in a high school which was in the Canterbury-Westland scheme as well as teaching mathematics at a night school which sat the external examination I was in the position to make such a comparison.

The second purpose of the research reported here was to compare the results of the night school class which sat the external examination with what they would have achieved if they had been in the internally assessed high school class taught by
the same teacher.

3.0 STUDY 1

3.1.0 Subjects

The school used in this study was a single sex boys' school located in a middle-class area in Christchurch.

3.2.0 School Organisation

3.2.1 Fifth Form

The school had been in the Canterbury-Westland School Certificate Scheme for the internal assessment of School Certificate Mathematics since 1975.

The scheme has two principal aspects:

1. Internal Assessment
2. The provision of courses at different levels to suit the ability of the pupils.

Internal Assessment Procedures

Mathematics was a compulsory subject for all Form five pupils and all of them sat the reference test. The first four weeks of the fifth form year were spent preparing for the test. Following this the course was broken up into topics, the order of which was selected to provide an easy transition from Form four work and to ensure that there was a reasonable spread of 'new' work throughout the year.

The testing programme was as follows:

Reference Test	(March)	75 minutes	10%
Term 1 Common Test	(early May)	50 minutes	15%
Term 2 Common Test	(late July)	2 hours	20%
Term 3 Common Test	(early November)	2 hours	20%
Individual Class Tests and Assignments (at least five)			35%

The 65 percent common element was used to establish comparability between classes. A pool of grades were then allocated to classroom teachers to award on the basis of the common and class tests.

The three common tests examined all the work covered for the year to that date. The students and their parents received a detailed description of the assessment programme and procedures at the beginning of the year.

Levels

At the fifth form level there were three distinct bands. There were three six-subject classes, five five-subject classes and one 'slow-learner' class. The two upper levels were streamed but not solely on the basis of mathematical ability. A single scheme covered both these levels and was the basis for the common tests. However within classes there were opportunities for extension and the top classes in each level were expected to cover more difficult aspects of the set topics. These were assessed in class tests and assignments. The 'slow learner' class used the Canterbury Mathematical Association's General Mathematics Certificate prescription as the basis of its course and all of the class entered for this award. Those pupils who also wished to obtain School Certificate sat the common tests to determine the grades to be allocated to the class.

3.2.2 Sixth Form Certificate

At sixth form level, mathematics was not compulsory. All of the sixth form mathematics classes were on at the same time and the classes were streamed using the final fifth form rankings from the previous year.

Internal Assessment Procedures

The Form 6 Mathematics course was taught in 6 units, approximately 2 units per term, as follows:

Unit 1 :	Coordinate Geometry; Algebra I
Unit 2 :	Sequences & Series; Algebra II
Unit 3 :	Functions & Graphs; Indices, Surds and Logarithms; Simultaneous Equations
Unit 4 :	Calculus
Unit 5 :	Probability & Statistics
Unit 6 :	Trigonometry

Testing was as follows

Common Test 1	Mid-March	(50 mins.)	15%
Common Test 2	Late June	(2 hours)	25%
Examination	Early November	(3 hours)	40%

All students sat the above tests which assessed all work for the year to that date.

In addition there was a 20% internal class component which consisted of at least two class tests or assignments each term.

The boys and their parents were advised of the assessment procedures at the beginning of the year.

A preliminary ranking of students was given after Common Test 2 and the final rankings which were sent to the Department of Education for the allocation of Sixth Form Certificate grades were available for inspection soon after the examination in November.

3.2.3 Seventh Form Mathematics with Statistics

All of the seventh form mathemtics classes were on at the same time and the classes were streamed using the final sixth form ranking from the previous year.

Assessment

The end-of-year 3 hour written external examination provided only 80% of the Bursary mark - the remaining 20% came from the two internally assessed practical projects the students presented during the year.

This was the second year of Mathematics with Statistics - and it differed markedly from all previous nationally recognised mathematics courses sat by the students in mathematics, in having a school assessed practical component contributing to the final results (7th form Geography has a similar system). The practical component offered a great deal to the student - the chance to research an area of mathematics knowing that time, encouragement and advice would be available. This 20% component of the Bursary mark was a reflection of ability in the practical areas of the syllabus not appropriately assessed in a 3 hour formal examination -for instance the ability to use computers effectively; to design, implement and analyse statistical projects or simulation experiments.

Success in this section did not need to have any relation to probable examination success - different attributes were being assessed.

Average/low examination results could be considerably enhanced by marks from quality practical work.

Midway through Term 1 a project common to all students was assigned with an end-of-term deadline (8%). Later in the term they nominated the topic which was to be their major individual effort (12%) - preliminary planning was marked no later than Mid-Term Break with the deadline for presentation of the final report being the second Monday in Term 3.

Full project details were provided to each student - topic ideas, teacher support, the marking schedule, the need for a planning/data book, the format of the final presentation; the calculator and computer requirement; individual work only...

In October the two project marks were combined (4% and 16%) and a summary mark list was forwarded to Wellington where scaling took place. The actual mark added to each students' examination mark was moderated by the school's performance in the Bursary examination and their position in the practical-work rank order.

3.2.4 The Handling of Marks for Internal Assessment

INTRODUCTION

The award of a grade not truly reflecting a student's performance with respect to his peers is not only unjust but could also have a significant effect on that student's future. Currently with many courses for national certificates being based in part or entirely on school-based marks valid assessment is particularly important.

For assessment to have any value two distinct areas must be addressed, regardless of subject area:

The Quality of the Testing - regularity of testing,
test content,
the marking schedule ...
- subjective factors which are the concerns of each teacher and her/his department: and

The Handling of the Marks - recording, weighting,
combining, interpreting...
- clinical factors that should be and can be validly managed in a near totally objective manner.

PARTICULAR

Testing

Each of the tests used for internal assessment in the school in this study, were set by a different teacher, moderated by another teacher, screened for face and content validity by all the mathematics teachers before the final design was decided on. The tests were considered to have met the following guidelines:

Guidelines for Checking Reliability

1. The test was long enough
2. The questions were clear
3. The time limits were realistic
4. The questions were of appropriate difficulty
5. The marking was effective
6. The instructions were clear
7. There was no choice of questions

Guidelines for Checking Validity

1. The questions were relevant, important
2. The topics were all assessed in appropriate proportions
3. All skills were assessed in appropriate proportions
4. The typing and presentation were adequate
5. All the students had an adequate opportunity to learn the material tested
6. Security was adequate to avoid cheating

Handling of Marks

The use of incorrect and/or inconsistent mark handling procedures can produce distortions in final rank order far inexcess of any that might be generated through the poor quality of testing!

Using the 'Motorised Mark Book' it is possible to employ straightforward acceptable routines to transform raw data to final marks in the knowledge that the same outcome will be obtained by others at other times with no more information than the raw data and the Assessment Statement.

Motorised Markbook

Motorised Markbook was written not to meet the needs of the individual teacher but rather to handle the statistics associated with combining common test scores from a number of classes. However a means has also been found which allows the classroom teacher to store and manipulate on a single disk the assessment data of up to 6 classes - printed results can range from the raw data through to a standardised weighted summary mark presented in alphabetic or rank order.

There are a number of <u>weaknesses</u> in the programme:

1. The limit of only 250 as the maximum mark reduces discrimination when combining a large number of tests: This can be overcome by use of the programme and a calculator.

2. There is very limited memory allocated for storing and handling the numbers appearing in the intermediate calculations - the associated rounding of results with resultant loss of accuracy cannot be easily corrected. The resultant effect is perhaps most apparent when the same results are combined with the only variable being the weighting factor (i.e. .1,.1,.3,.4 gives a differing final rank order than the equivalent weighting of 1,1,3,4, or 2,2,6,8). All that can be done is to decide at the beginning of the year on how weightings are to be expressed and then be consistent throughout the year- there may be slight injustices. However these are likely to be insignificant in comparison to the injustices that arose from previous methods.

3.3.0 Sample

There were two samples used in this study. The first was composed of those students who sat School Certificate Mathematics in 1986, Sixth Form Certificate in 1987 and Bursary statistics in 1988. The second sample sat School Certificate in 1987, Sixth Form Certificate in 1988 and Bursary in 1989.

3.4.0 Material

We obtained the following information from each student's cumulative record for mathematics as it became available:

Sample 1

1986 School Certificate Reference Test Mark and Ranking

1986 Final fifth form ranking

1987 Final sixth form ranking

1988 a. Final Bursary mark in Statistics (internally assessed component and exam mark combined)

b. Mark from internally assessed component.

c. Mark from the Bursary exam.

Sample 2

1987 School Certificate Reference Test Mark and Ranking.

1987 Final fifth form ranking

1988 Final sixth form ranking

1989 Final Bursary mark in Statistics. (the breakdown of marks into the internally assessed component and examination mark was not available at the time of writing.

3.5.0 Design and Procedure

We calculated means and standard deviations for the School Certificate Reference Test and Bursary marks for both samples and the school population. In a study of this nature, one inevitably encounters missing data. Sometimes test scores or other information was missing because students in the sample were either not at the school for all of the three years, or did not continue with mathematics beyond the fifth or sixth form. Students with missing scores were included in the statistical tests where those scores were not involved. The specific number of students used in each comparison is reported below. To ensure the final sample was representative of the schools' results a hypothesis test was carried out on the diffences of the sample and population means for the Reference Tests and Bursary marks.

The analysis of the data was largely performed using the features of the computer spreadsheet "SuperCalc 5". The following correlation analyses were performed:

Sample 1

1. 1986 School Certificate Reference Test ranks

with
a. 1986 final Fifth Form rankings
b. 1987 final Sixth Form rankings

2. 1986 Reference Test marks

with
a. 1988 final Bursary marks
b. 1988 internally assessed Bursary marks
c. 1988 Bursary examination marks

Sample 2

1. 1987 reference test

   with a. 1987 Final Fifth form rankings

   b. 1988 final Sixth Form rankings

2. 1987 final Fifth Form rankings

   with 1988 final Sixth From rankings

3. 1987 Reference Test marks

   with 1989 final Bursary marks

Where information was not available in the form of numberical values - the Spearman Rank Order Correlation Coefficients were calculated and tested for significance. Where numerical values were available for both sets of data Pearson Correlation Coefficients were calculated and tested for significance.

Scatter diagrams and regression lines were also produced for each comparison.

3.6.0 Results

3.6.1 Descriptive Results

Table 12 contains the descriptive results of all variables included in this study.
The population figures refer to the statistics for all the pupils in the experimental school who sat the particular test. The sample figures refer to the statistics for those

students who were included in the final comparisons of the Reference Tests and Bursary Statistics.

Descriptive Statistics for All Variables Used in the Study

	Population			Sample Group			t
	N	M	SD	N	M	SD	
1986 S.C. Reference Test	216	68.22	21.60	97	81.21	10.90	*** 11.63
1987 S.C. Reference Test	239	63.57	19.83	107	75.03	13.07	*** 8.92
1988 Final Bursary Mark	105	60.50	15.27	97	62.68	15.14	* 1.42
1988 Bursary Examination Mark (80%)	105	61.30	15.75	97	63.03	15.97	1.06
1988 Bursary Internal Assessment Mark adj (20%)	105	61.91	15.14	97	63.74	16.59	* 1.83
1989 Final Bursary Mark	120	58.12	13.47	107	57.64	13.45	-0.37

* $p < 0.10$
** $p < 0.05$
*** $p < 0.001$

Table 12

3.6.2 Regression Results

The results of the regressions of the dependent variables on the independent variables are shown in Tables 13 - 16 and graphs 1 -9. The tables report the following information: number in the sample N, regression coefficient (B), standard error of B, correlation coefficient R and the determination coefficient R^2.

Results of the Regression Analysis of the 1986 School Certificate Reference Test on the Dependent Variables

Dependent Variables	N	B	Error B	R**	R^2
Final Fifth Form Ranking 1986	239	0.92	0.03	0.92	0.84
Final Sixth Form Ranking 1987	163	0.98	0.03	0.69	0.47
Final Bursary Statistics Mark 1988	97	0.88	0.12	0.64	0.41
Internally Assessed Bursary Statistics Mark 1988 (20%)	97	0.42	0.17	0.28	0.08
Bursary Statistics Examination Mark 1988 (80%)	97	0.94	0.13	0.64	0.42

Table 13

Regression Analysis for the 1988 Bursary Exam Marks and Bursary Internal Assessment

N	B	Error B	R**	R^2
97	0.55	0.10	0.54	0.29

Table 14

Results of the Regression Analysis of the 1987 School Certificate Reference Test on the Dependent Variables

Dependent Variables	N	B	Error B	R**	R^2
Final Fifth Form Ranking 1987	251	0.86	0.03	0.90	0.80
Final Sixth Form Ranking 1988	176	0.82	0.05	0.61	0.39
Final Bursary Statistics Mark 1989	107	1.13	0.15	0.60	0.36

Table 15

Regression Analysis for the 1987 Final Fifth Form Rankings and 1988 Final Sixth Form Rankings

N	B	Error B	R**	R^2
176	1.02	0.13	0.79	0.62

Table 16

(Note: For tables 13 - 16

** The level of significance was determined by using the statistical tables for the 'Critical values for correlation coefficients' in Rohif and Sokal. $p > 0.01$ in all cases)

Spearman Rank Order Correlations computed for the final Fifth Form, final Sixth Form placings and the Reference Test showed a significant correlation for each pair of variables at the 0.01 level for both samples. (Tables 13 & 15) The correlation between the 1987 final Fifth Form rankings and 1988 Final Sixth form rankings was also significant at the 0.01 level. (Table 14)

The strongest correlations were those between the Reference Tests and the Fifth Form ranks for the same year. These were followed by the 1987 Fifth Form and 1988 Sixth Form ranks. Finally the Reference Tests and the Sixth Form rankings the following year had the lowest correlations. These results were to be expected. The Reference Test and the Fifth Form results reflected that they were the closest together time-wise and also the 10% component of the final Fifth Form results that came from the Reference Test. In the other comparisons the strength of the relationship diminished as the time between the gradings increased. Thus the relationship between the 1987 final Fifth Form and 1988 Sixth Form was stronger than that between the 1987 Reference Test and 1988 Sixth Form.

The Pearson Product Moment Correlation between the Reference Test marks and the final Bursary marks was positive and significant for both samples; sample 1, $r = 0.64$. $p > 0.01$, sample 2, $r = 0.60$, $p > 0.01$. The correlations indicate that there was a relationship between the Reference Test marks and the final Bursary marks; the reference test accounting for between 36 to 41% of the variance in the Bursary marks. When the final 1988 Bursary mark was broken into its internal and externally assessed components the correlations between the Reference Test and both components in turn were also significant at the 0.01 level. However the Reference Test accounted for only 8% of the variance in the internally assessed component as compared with 41% in the external exam.

4.0 STUDY 2

4.1.0 Subjects

The subjects used in this study were students who had enrolled in a night school class in February 1988 with the intention of sitting school certificate mathematics by external examination in November of the same year. The comparison group consisted of the fifth form described in study 1 for whom mathematics was compulsory and internally assessed.

4.2.0 Procedure

The same teacher taught the night school students and one fifth form internally assessed class for the whole year. The same material was taught to the night school students on a Tuesday night from 6 to 9 as was covered in four fifty minute periods in the high school during the same week. Both classes followed the same testing programme used for internal assessment within the high school. However the night school class was told that, whilst the assessment was important it would only be used towards their final grade if they needed to apply for an aegrotat if for some reason they were unable to sit the external examination.

After the high school rankings were computed the equivalent data for the night school was added to the same computer file and the night school students were assigned theoretical grades on the basis of their year's work relative to the pupils in the school. These grades were then compared to the grades they obtained from the external examination.

4.3.0 Results

Night School Results

Student	Reference Test Mark	"Internally Assessed Grade"	Examination Mark	Final Grade
A	36	A2	74	A2
B	40	B1	57	B1
C	32	B1	57	B1
D	41	B2**	56	B1
E	25	B2	54	B2
F	22	B1*	52	B2
G	21	B2	49	B2
H	20	B2	48	B2
I	27	B2*	44	C1
J	25	C1	41	C1
K	28	C1	41	C1
L	17	B2*	38	C1
M	17	C1	38	C1
N	16	C1*	29	C2
O	10	C2	21	C2

* Higher Internally Assessed Grade

** Lower Internally Assessed Grade

Although the night school class originally had 30 students only 15 students took the external examination - the rest dropped out for various reasons. Ten of these students achieved the same grade in the external examination they would have been assigned using internal assessment. Whilst four students obtained lower grades on the external assessment they were only by one grade. Closer examination of their final marks and internal assessment rankings revealed they were very close to the borderline in both cases. A few more "marks" in the external examination would have meant a higher grade. Alternatively a few places lower in the rankings would have meant a lower grade in internal assessment.

Only one student obtained a higher grade in the external examination. Again he was close to the borderline for the lower grade in the external examination and just below the cut-off for the higher grade in the internal assessment.

5.0 DISCUSSION

Table 17 shows the descriptive statistics for the 1986 and 1987 reference tests for the pupils in the study and all of the candidates in the Canterbury Westland Scheme.

Descriptive Statistics for the Reference Tests

Reference Test	School in Study			All Schools in Scheme			
	N	M	SD.	N	M	SD.	t
1986	216	68.22	21.60	6093	47.00	22.80	14.16*
1987	239	63.57	19.83	6347	45.67	20.83	13.67*

* $p < 0.01$

Table 17

Clearly the schools' results were statistically different from the combined results of all of the schools in the scheme. This was to be expected because the sample was limited to one school enrolling middle class boys. Hence any generalisations drawn from this investigation should be considered with caution.

The major concern of this investigation was the School Certificate Reference Test and the implications of these results on the fifth form and sixth form grades. In particular the study investigated whether the School Certificate Reference Test accurately predicted a) final fifth form rankings and b) final sixth form rankings.

The analysis showed with 99% confidence that there were positive linear correlations between the 1986 and 1987 reference tests and each of the dependent variables in turn.

The relationships between the Reference Tests and final fifth form rankings were the strongest with 84 and 80% of the variance in the fifth form rankings accounted for by a linear relationship with the Reference Test. It should be remembered that 10% of the final fifth form mark came from the Reference Test. Even taking this factor into account the Reference Test at the beginning of the year appeared to accurately predict the students' performance throughout the year. Changes of attitude and academic development of students during the fifth form year either were not significant in this school or the reference test could be said to have catered for them.

Whilst Study 1 showed a high correlation between the reference test rankings and the final fifth form rankings this did not answer the question raised by the non-participating schools about the distribution of grades allocated to the schools. This was the purpose of the second study. Because this study was limited to a single night school class and the final numbers were statistically too small, no generalisations could be made from the results. However, ignoring these limitations and other pros and cons for internal assessment, the study could be considered to have favoured internal assessment for 27%, neither favoured nor disfavoured internal assessment for 67% and disfavoured internal assessment for 0.07% of the students.

In addition, the night school class was not representative of the high school sample. There were eight females in the final night school class and their ages varied from 17 to 40. The high school was an all boys school and their ages varied from 15 to 17. In addition the high schools' distribution of grades was skewed to the right in the comparison with all the schools in the scheme. Hence the night school students might have also been awarded the lower grade for a similar achievement in another interally assessed school more representative of the total population of the scheme. However five years experience of teaching to both the internal and external examinations lead the author to feel the students possibly deserved the higher grades. In particular, the students had been observed over a longer time period, in a greater range of situations and had undergone many more assessments. More important, though, was the evidence of additional stress experienced by these students as the examination approached and previous failures were remembered. After the initial dropout, as students discovered they were overcommitted, the next main dropout was in the third term. The one student who improved his grade had obtained a final internal assessment grade which was lower than expected from his reference test mark and he noticeably increased his effort as he approached the examination. This does not provide support for the external examination. An even distribution of effort throughout the year encouraged by internal assessment rather than a last minute cram is a better preparation for what is expected in adult life.

As a result of the study, this year, this particular School Certificate Mathematics night school class, is being offered the chance to sit for the Canterbury Mathematical Association Certificate, as well as School Certificate. It is hoped that the more frequent feedback and success offered by the certificate will cut down the number of dropouts and provide the students with the confidence they need to do themselves justice in the final examination.

They certainly do not lack in motivation. At the very worst they will have a certificate stating what they have mastered and at the best they will have School Certificate Mathematics - which because of previous experience is very important to them. Internal assessment is not an option available to the night school, these students who have decided to give mathematics a second try deserve better than to again experience failure because of a one-off examination. The lowest correlations were obtained for the comparison of the reference test with the sixth form rankings where 47% and 39% of the variation in the sixth form rankings could be accounted for by the 1986 and 1987 Reference Tests respectively. There were a number of factors that would have influenced these correlations.

Firstly, the two years between the two assessments were certainly a long enough period of time for maturation and history to have affected the outcomes. This was further evidenced by the decrease in correlations observed between the Reference Test and fifth form rankings; the fifth form rankings and sixth form rankings; and the Reference Test and sixth form rankings respectively.

Using the rank correlations had a useful advantage over the Pearson product moment correlations in that extreme observations in the rank correlation tests will never produce a large rank difference.

Mortality was also a serious threat in this study because 30% of the students who took the Reference Test were no longer taking mathematics at the school at the time the sixth form rankings were obtained.

The correlations between the Reference Test and Bursary results were carried out to further test how well the Reference Test marks predicted success in later education. Because both sets of data were available as scores, Pearson product moment correlations were calculated. The determination coefficients showed that 39 - 41% of the variance in the Bursary marks could be explained by the Reference Test. Again history, maturation and mortality would have affected these outcomes.

The mean of the Reference Test for the sample who sat Bursary was significantly different from the mean of the original fifth form population (Table 12) and only 40% of the original fifth form actually sat Bursary Statistics. These analyses were further complicated because the final Bursary mark was standardised and made up of the internally assessed component (20%) and the final examination mark (80%). Table 13 shows that the correlation between the internally assessed Bursary component and the Reference Test was only 0.28 ($R^2 = 0.08$).

In addition, the Reference Test consisted of 60 one mark questions and took 75 minutes, whereas the Bursary examination was a three hour paper and tested more analytical skills. Hence there were several very extreme observations in the Bursary data and those alone could mean that the correlation coefficients calculated may not be a good description of the association that existed between the two variables.

This year 50% in the Bursary examination was used as the cut-off score for entry into Stage I statistics at Canterbury University. Because of this, limited entry to a subject which is a pre-requisite for a number of different courses, a predictor of a student's performance in the Bursary examination may be of use to those students who are considering career choices which need Stage I statistics.

The results of the regression analysis here indicates that the reference test could be used to predict a student's future progress in the subject. In particular, those students who are clearly performing below expectaions, might benefit from such a personal discussion. Early intervention may prevent possible restriction to choices of career in the future.

6.0 CONCLUSION

Of the 56,852 students in their first year in form five on 1st March 1988, 11049 (19%) left school during, or at the end of the year. Of the 11049 students, 5148 (48%) left school with no School Certificate qualification.[2] It was clear that those who needed a nationally recognised leaver qualification did not get one, while the vast majority of those who did well in School Certificate remained at school for a form six year.

At the moment the external School Certificate examinations appear to fill only one function: they are used to moderate school-based assessments for Sixth Form Certificate. Once Sixth Form Certificate becomes an achievement-based assessment award, as proposed by the Board of Studies,[2] they present School Certificate will become redundant. Some will argue that the issue is about standards. However School Certificate is a norm-referenced award.

[2] Board of Studies Newsletter. No5 August 1989

In norm - referenced examinations the award of grades is heavily influenced by the relative overall performance of candidates and the standards of performance required for particular grades are based on assumptions about the proportions of candidates likely to gain the grades. Any notion of standards in School Certificate exists only in the minds of those who understand statistical moderation (commonly known as scaling).

The problem of determining standards in a way that applies across schools and subject areas is one that has worried educators, parents, employers and the community generally. The present solution - and an increasingly unpopular one today-has been to peg standards to a common external examination, but that creates other problems. Most notably, external examinations encourage stress and reproductive learning on the part of the students, while teaching to the external syllabus with a view to maximising marks strongly encourages a piecemeal packaging of the syllabus components, leading teacher and student alike towards a multistructural conception of learning.

The Commonwealth Schools Commission (1987) points out the numerous disadvantages of external examinations, indicating that 'the strengths of school-based assessments are, in general, the mirror image of the weakness of external examinations'. The wide range of abilities and achievements of the candidates, make it extremely difficult to set questions so that all candidates have adequate opportunities to show what they know, understand and can do.

The fact that the Canterbury-Westland Scheme for the internal assessment of mathematics is now entering its 15th year, must be testimony to its validity in assessing and grading students' mathematical ability at fifth form level. The moderation procedures used in the scheme are designed so that students at any level are not advantaged or disadvantaged.

Despite its limitations, this study of the scheme at one middle class, boys high school, could be said to provide further evidence to support the scheme.

Statistical Tables

F. James Rohlf
Robert R. Sokal

State University of New York
at Stony Brook

W. H. Freeman and Company
San Francisco

TABLE Y. Critical values for correlation coefficients.

This table furnishes 0.05 and 0.01 critical values for product-moment correlation coefficients r and multiple correlation coefficients R. These values are given for every degree of freedom between $\nu = 1$ and $\nu = 30$ and selected degrees of freedom between $\nu = 30$ and $\nu = 1000$. This table is used to test the null hypothesis that the correlation coefficient of the population from which the sample has been taken is zero. Under such conditions the following t-test for significance of the product-moment correlation coefficient r applies:

$$t = r/\sqrt{(1 - r^2)/(n - 2)} = r\sqrt{(n - 2)/(1 - r^2)} \qquad \nu = n - 2$$

where n is the sample size (number of pairs of variates). The critical values in the table are computed by entering the correct values of $t_{\alpha[\nu]}$ and n in the above equation and solving for r. The critical values for the multiple correlation coefficients are based on a different formula involving the F-distribution.

To test the significance of a correlation coefficient, the sample size n upon which it is based must be known. Enter the table for $\nu = n - 2$ degrees of freedom and consult the first column of values headed "independent variables." For example, for a sample size of $n = 28$ and $\nu = 28 - 2 = 26$, the critical values of r are found to be 0.374 at the 5% level and 0.478 at the 1% level. Thus, for an observed correlation coefficient $r = 0.31$ in a sample of 28 paired observations, one would be led to conclude that the correlation between the variables concerned is not significantly different from zero. Negative correlations are considered as positive for purposes of this test. More details on significance testing of correlation coefficients are given in Section 15.4 and Box 15.3.

The other three columns in the table give critical values for a multiple correlation coefficient involving 2, 3, and 4 independent variables. Degrees of freedom for such a problem are $\nu = n - m$, where n is the sample size and m is the number of variables, both dependent and independent. Thus, for a sample value of $R = 0.42$ based on a sample of 50 items and measurements of four variables (one dependent plus three independent), one would conclude that it is significant at $P = 0.05$ but not at $P = 0.01$. The appropriate degrees of freedom are $50 - 4 = 46$, requiring interpolation, but since these conclusions are true for both $\nu = 45$ and $\nu = 50$, bracketing the correct value for the degrees of freedom, one need not interpolate.

This table is reproduced by permission from *Statistical Methods*, 5th edition, by George W. Snedecor, © 1956, by The Iowa State University Press.

TABLE Y. Critical values for correlation coefficients.

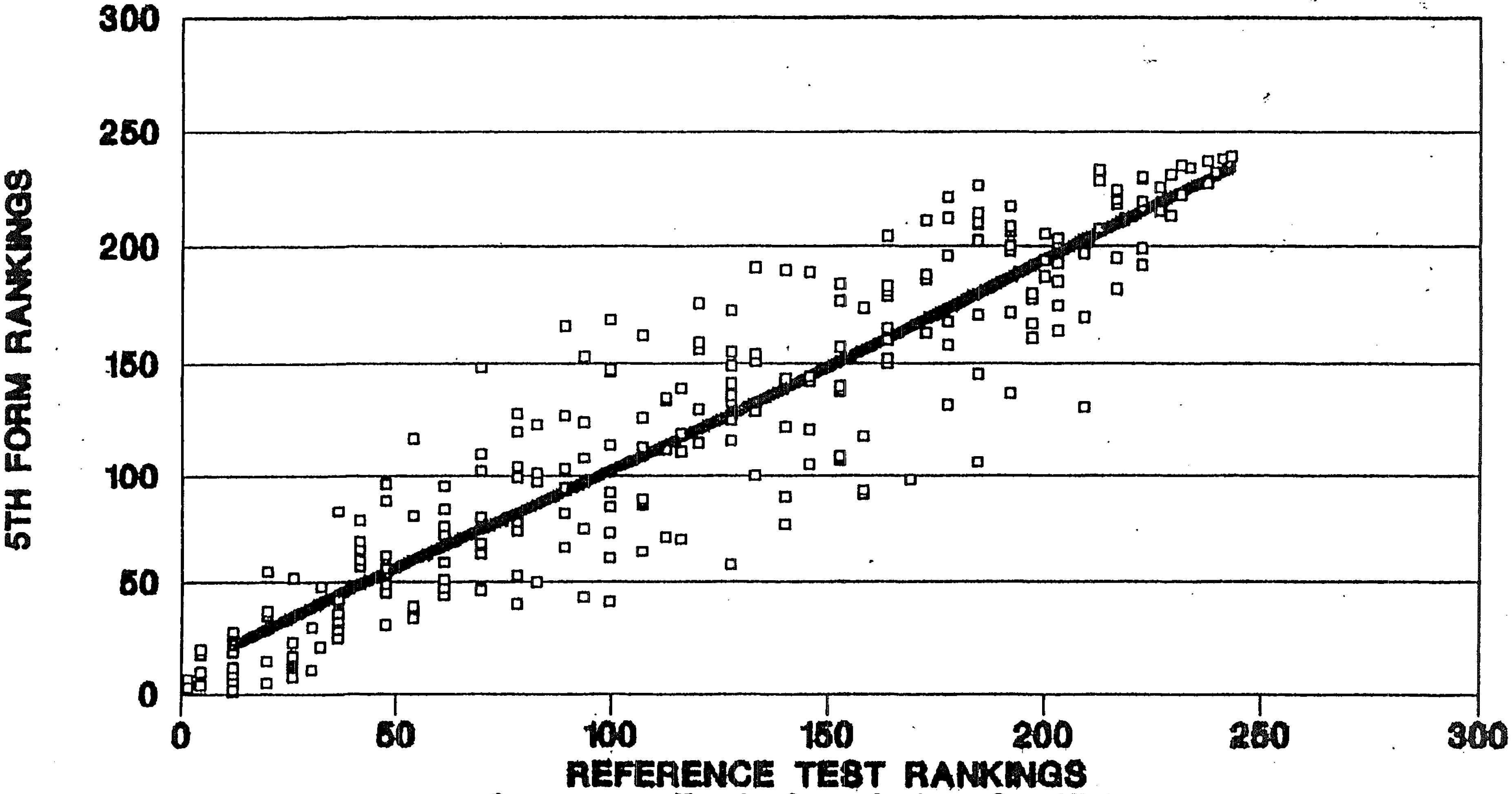
		Number of independent variables						Number of independent variables			
ν	α	1	2	3	4	ν	α	1	2	3	4
1	.05	.997	.999	.999	.999	16	.05	.468	.559	.615	.655
	.01	1.000	1.000	1.000	1.000		.01	.590	.662	.706	.738
2	.05	.950	.975	.983	.987	17	.05	.456	.545	.601	.641
	.01	.990	.995	.997	.998		.01	.575	.647	.691	.724
3	.05	.878	.930	.950	.961	18	.05	.444	.532	.587	.628
	.01	.959	.976	.983	.987		.01	.561	.633	.678	.710
4	.05	.811	.881	.912	.930	19	.05	.433	.520	.575	.615
	.01	.917	.949	.962	.970		.01	.549	.620	.665	.698
5	.05	.754	.836	.874	.898	20	.05	.423	.509	.563	.604
	.01	.874	.917	.937	.949		.01	.537	.608	.652	.685
6	.05	.707	.795	.839	.867	21	.05	.413	.498	.522	.592
	.01	.834	.886	.911	.927		.01	.526	.596	.641	.674
7	.05	.666	.758	.807	.838	22	.05	.404	.488	.542	.582
	.01	.798	.855	.885	.904		.01	.515	.585	.630	.663
8	.05	.632	.726	.777	.811	23	.05	.396	.479	.532	.572
	.01	.765	.827	.860	.882		.01	.505	.574	.619	.652
9	.05	.602	.697	.750	.786	24	.05	.388	.470	.523	.562
	.01	.735	.800	.836	.861		.01	.496	.565	.609	.642
10	.05	.576	.671	.726	.763	25	.05	.381	.462	.514	.553
	.01	.708	.776	.814	.840		.01	.487	.555	.600	.633
11	.05	.553	.648	.703	.741	26	.05	.374	.454	.506	.545
	.01	.684	.753	.793	.821		.01	.478	.546	.590	.624
12	.05	.532	.627	.683	.722	27	.05	.367	.446	.498	.536
	.01	.661	.732	.773	.802		.01	.470	.538	.582	.615
13	.05	.514	.608	.664	.703	28	.05	.361	.439	.490	.529
	.01	.641	.712	.755	.785		.01	.463	.530	.573	.606
14	.05	.497	.590	.646	.686	29	.05	.355	.432	.482	.521
	.01	.623	.694	.737	.768		.01	.456	.522	.565	.598
15	.05	.482	.574	.630	.670	30	.05	.349	.426	.476	.514
	.01	.606	.677	.721	.752		.01	.449	.514	.558	.591

ν	α	Number of independent variables 1	2	3	4
35	.05	.325	.397	.445	.482
	.01	.418	.481	.523	.556
40	.05	.304	.373	.419	.455
	.01	.393	.454	.494	.526
45	.05	.288	.353	.397	.432
	.01	.372	.430	.470	.501
50	.05	.273	.336	.379	.412
	.01	.354	.410	.449	.479
60	.05	.250	.308	.348	.380
	.01	.325	.377	.414	.442
70	.05	.232	.286	.324	.354
	.01	.302	.351	.386	.413
80	.05	.217	.269	.304	.332
	.01	.283	.330	.362	.389
90	.05	.205	.254	.288	.315
	.01	.267	.312	.343	.368
100	.05	.195	.241	.274	.300
	.01	.254	.297	.327	.351
125	.05	.174	.216	.246	.269
	.01	.228	.266	.294	.316
150	.05	.159	.198	.225	.247
	.01	.208	.244	.270	.290
200	.05	.138	.172	.196	.215
	.01	.181	.212	.234	.253
300	.05	.113	.141	.160	.176
	.01	.148	.174	.192	.208
400	.05	.098	.122	.139	.153
	.01	.128	.151	.167	.180
500	.05	.088	.109	.124	.137
	.01	.115	.135	.150	.162
1,000	.05	.062	.077	.088	.097
	.01	.081	.096	.106	.115

APPENDIX

Graph 1

Correlation Analysis for Rankings in S.C.Ref.Test and 5th Form Year End 1986



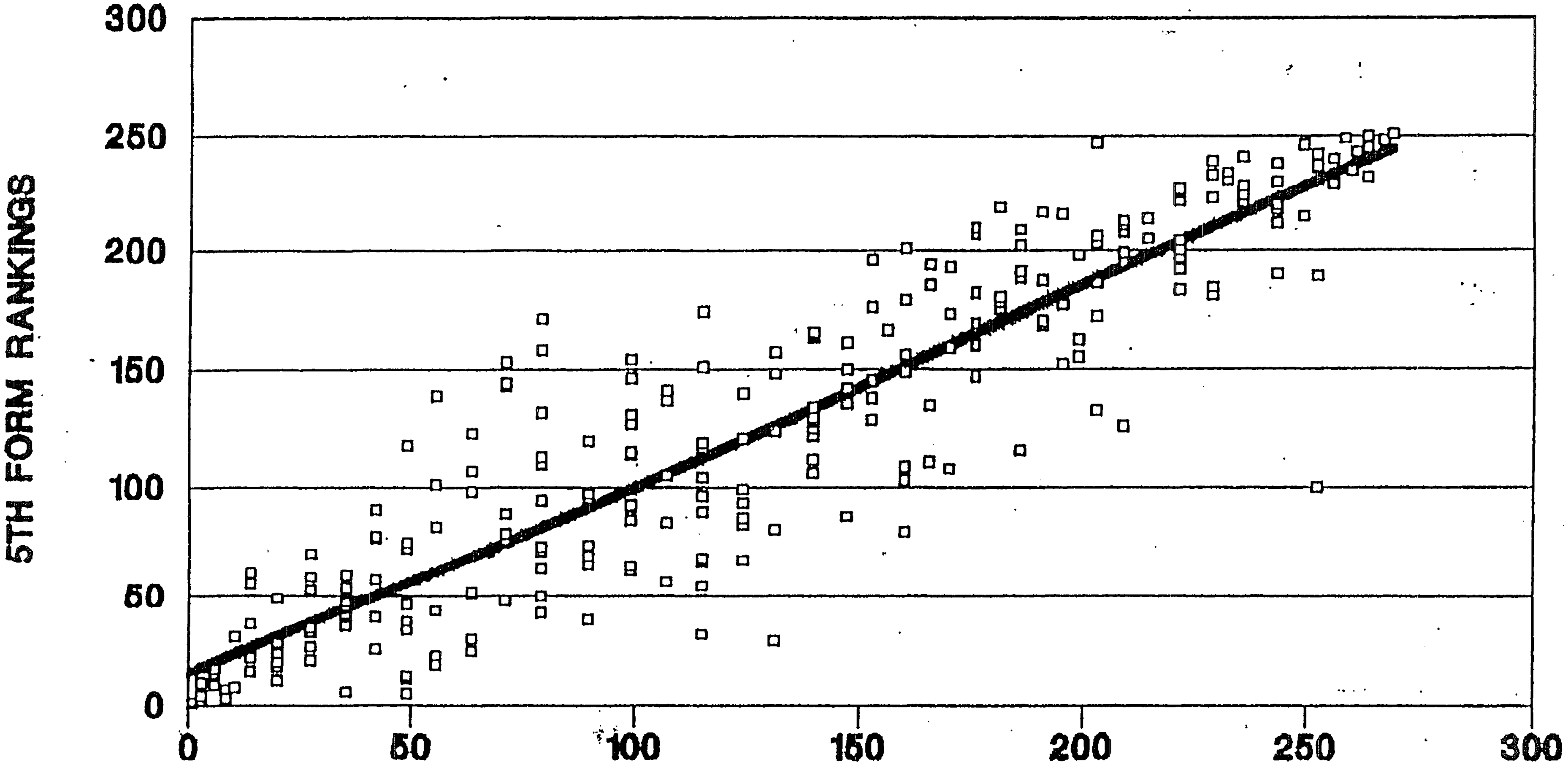
Spearman Rank Correlation Coefficient

Rs = 0.92

Regression Line At Y = 0.92X + 11.17

Graph 2

Correlation Analysis for Rankings in S.C.Ref.Test and 5th Form Year End 1987



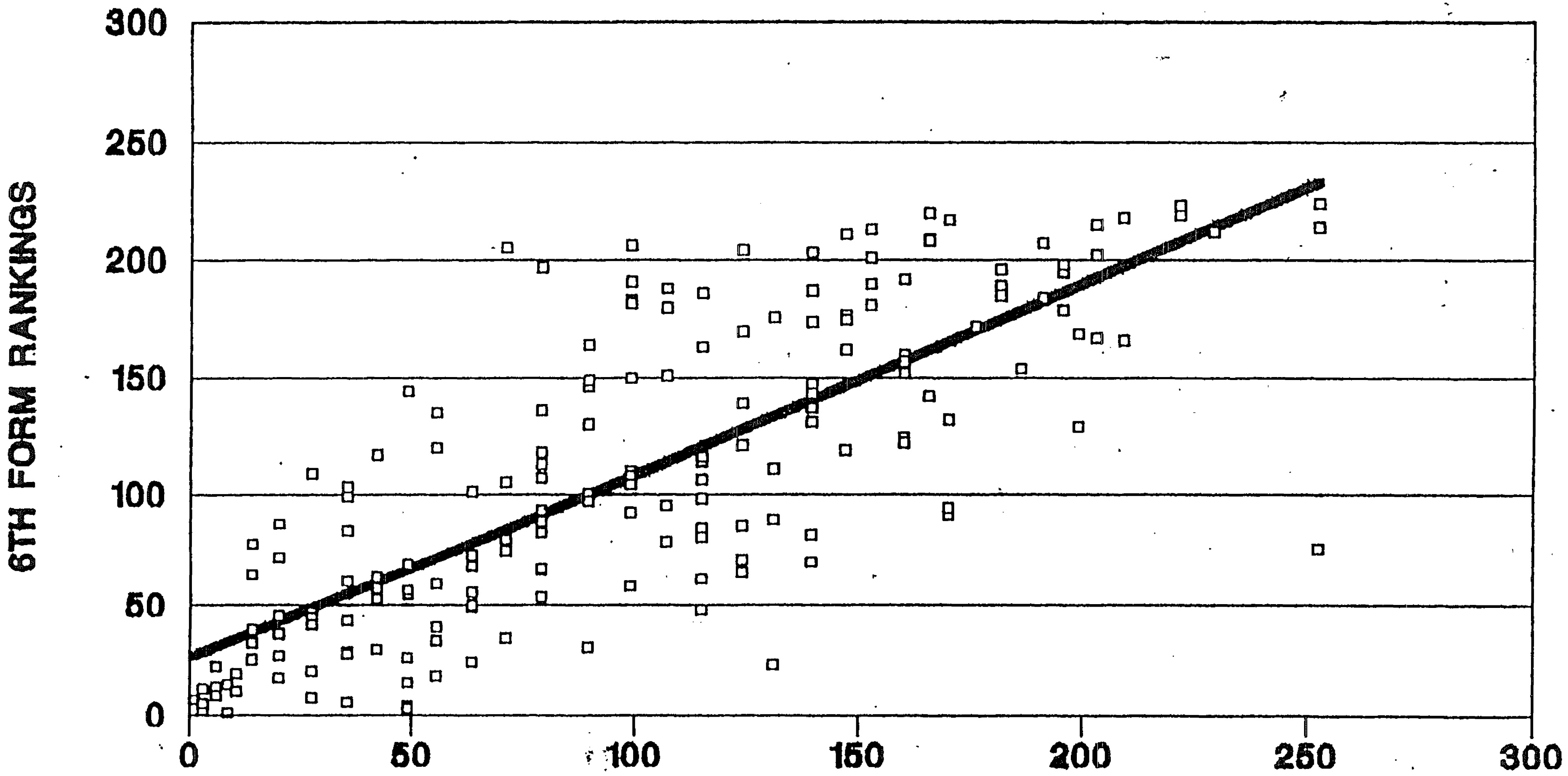
Spearman Rank Correlation Coefficient

Rs = 0.90

Regression Line At Y = 0.86X + 14.51

Graph 3

Correlation Analysis for Rankings in S.C. Ref. Test 1987 and 6th Form 1988



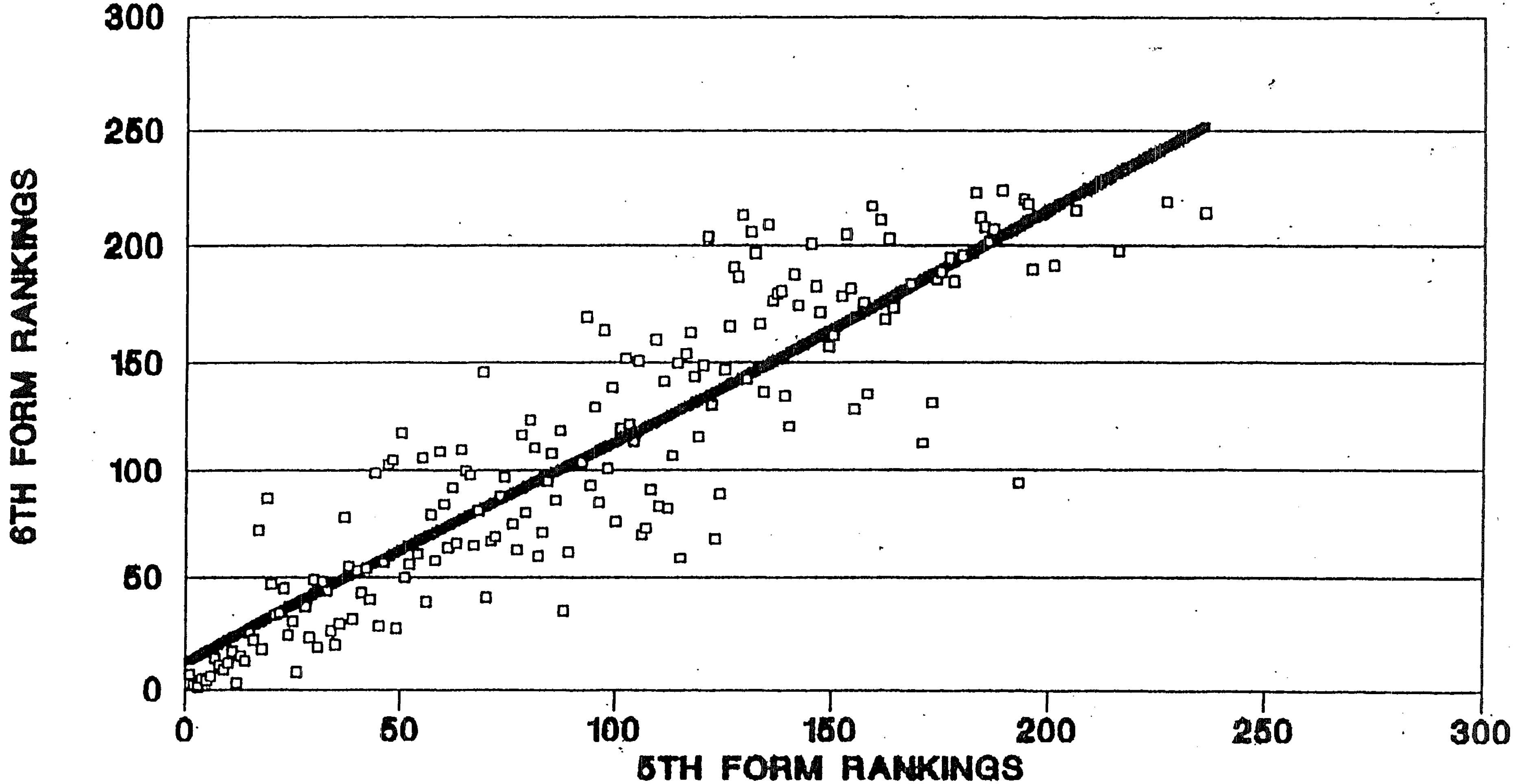
Spearman Rank Correlation Coefficient

Rs = 0.61

Regression Line At Y = 0.82X + 26.68

Graph 4

Correlation Analysis for Rankings in 5th Form 1987 and 6th Form 1988

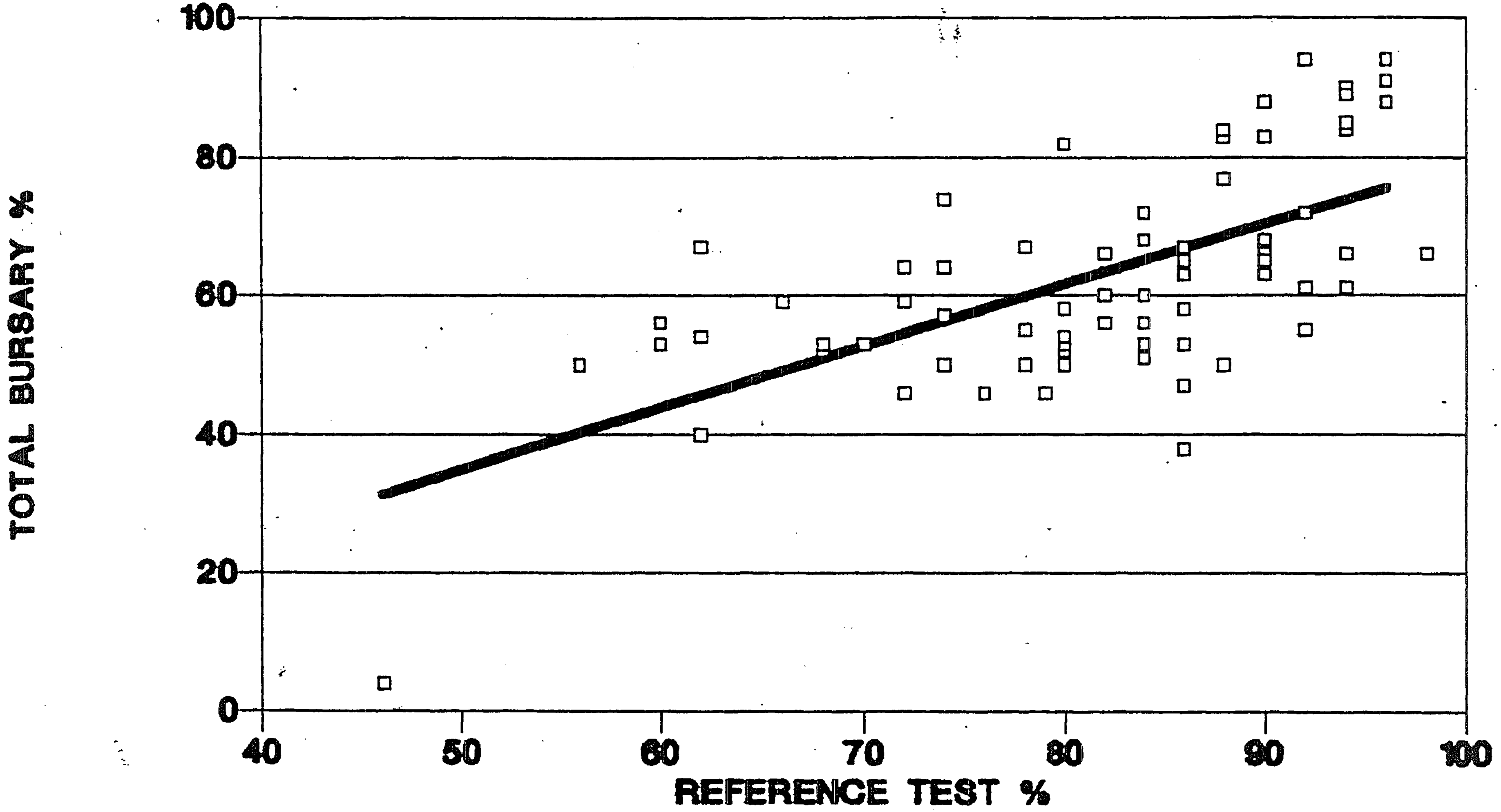


Spearman Rank Correlation Coefficient
Rs = 0.79
Regression Line At Y = 1.02X + 12.33

Graph 5

Reference Test % versus Total Bursary %

(1986) (1988)



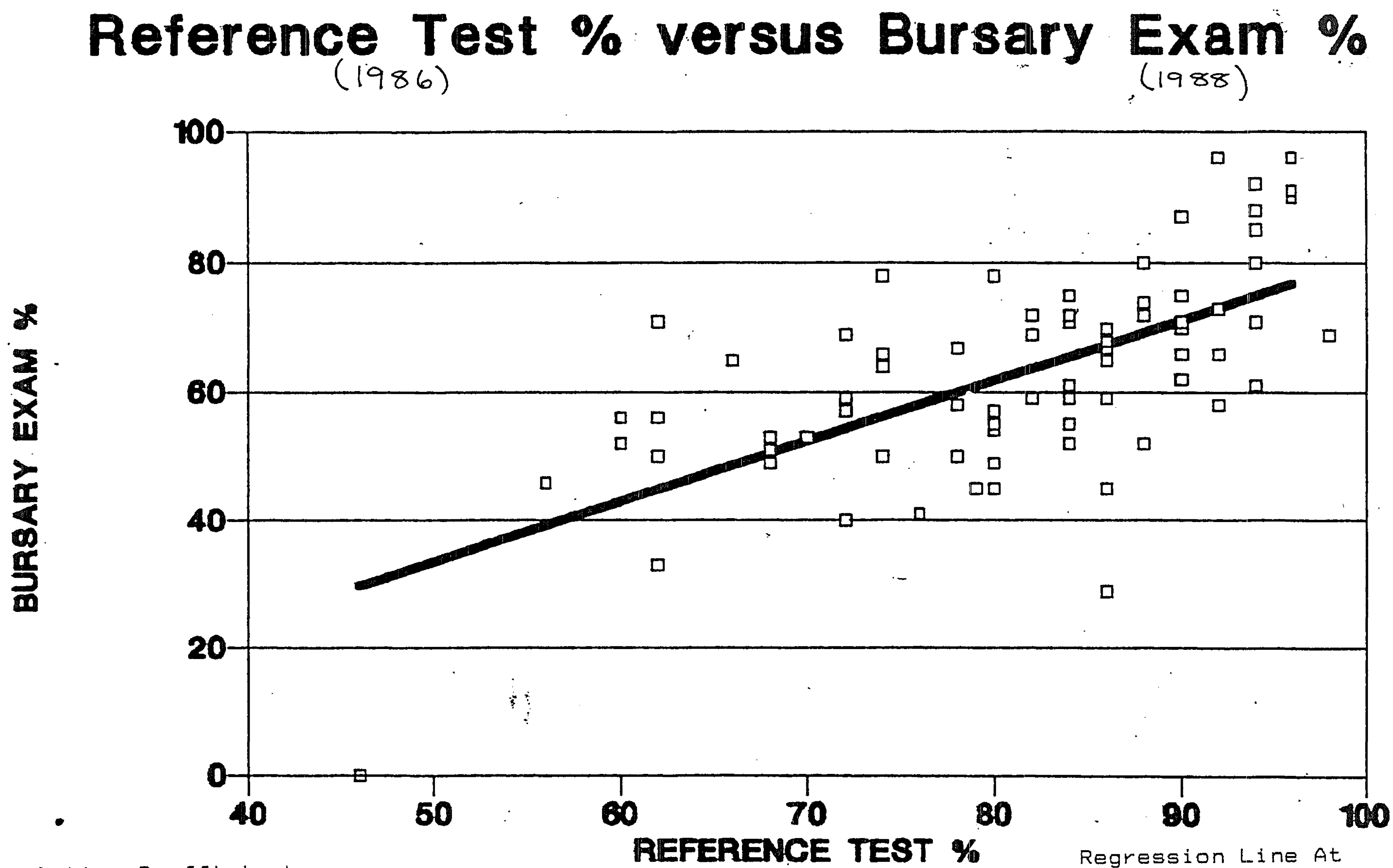
Pearson Correlation Coefficient

R = 0.64

Regression Line At

$Y = 0.88X - 9.07$

Graph 6



Pearson Correlation Coefficient

R = 0.64

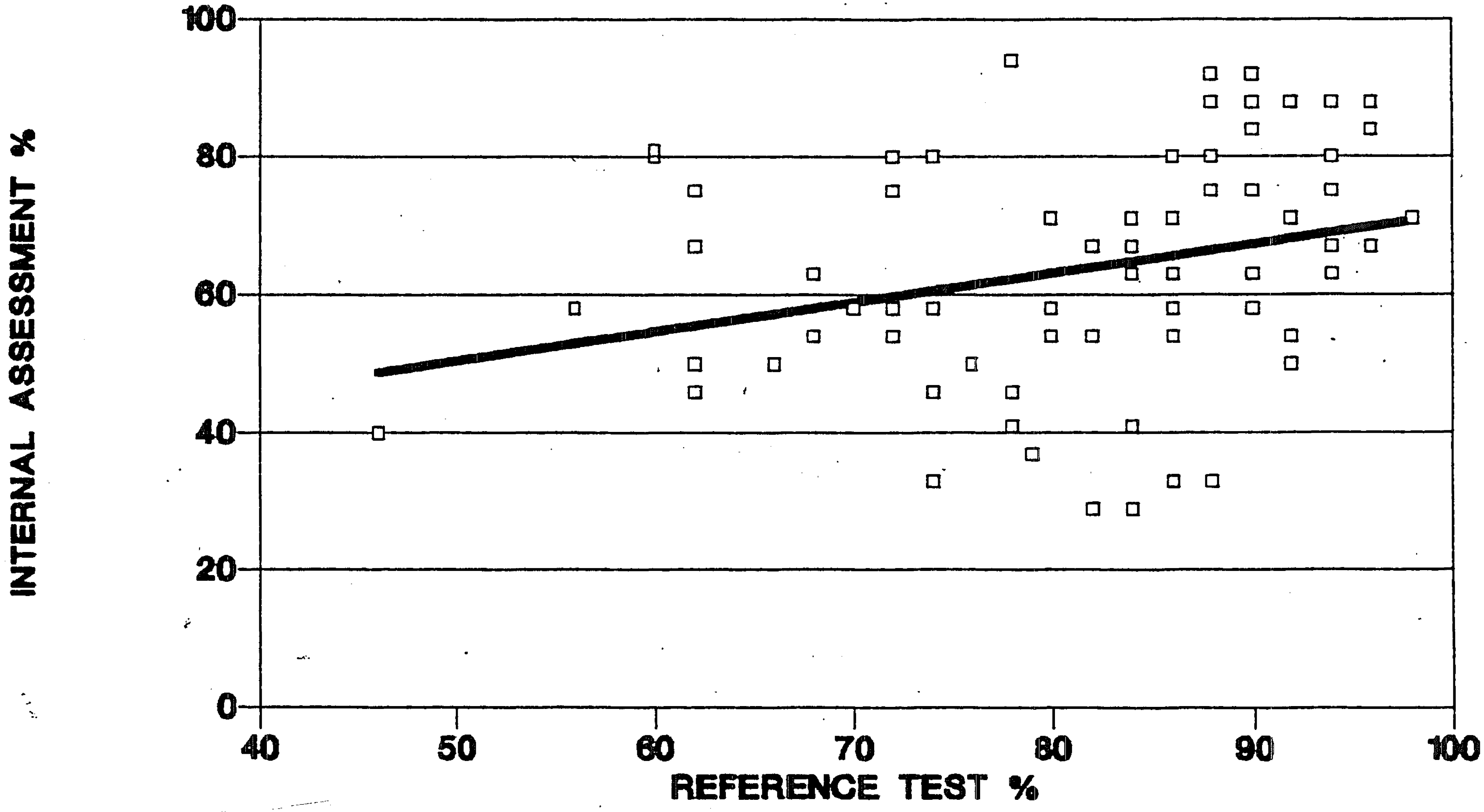
Regression Line At

Y = 0.94X - 13.36

Graph 7

Reference Test % versus Bursary Internal %

(1986) (1988)



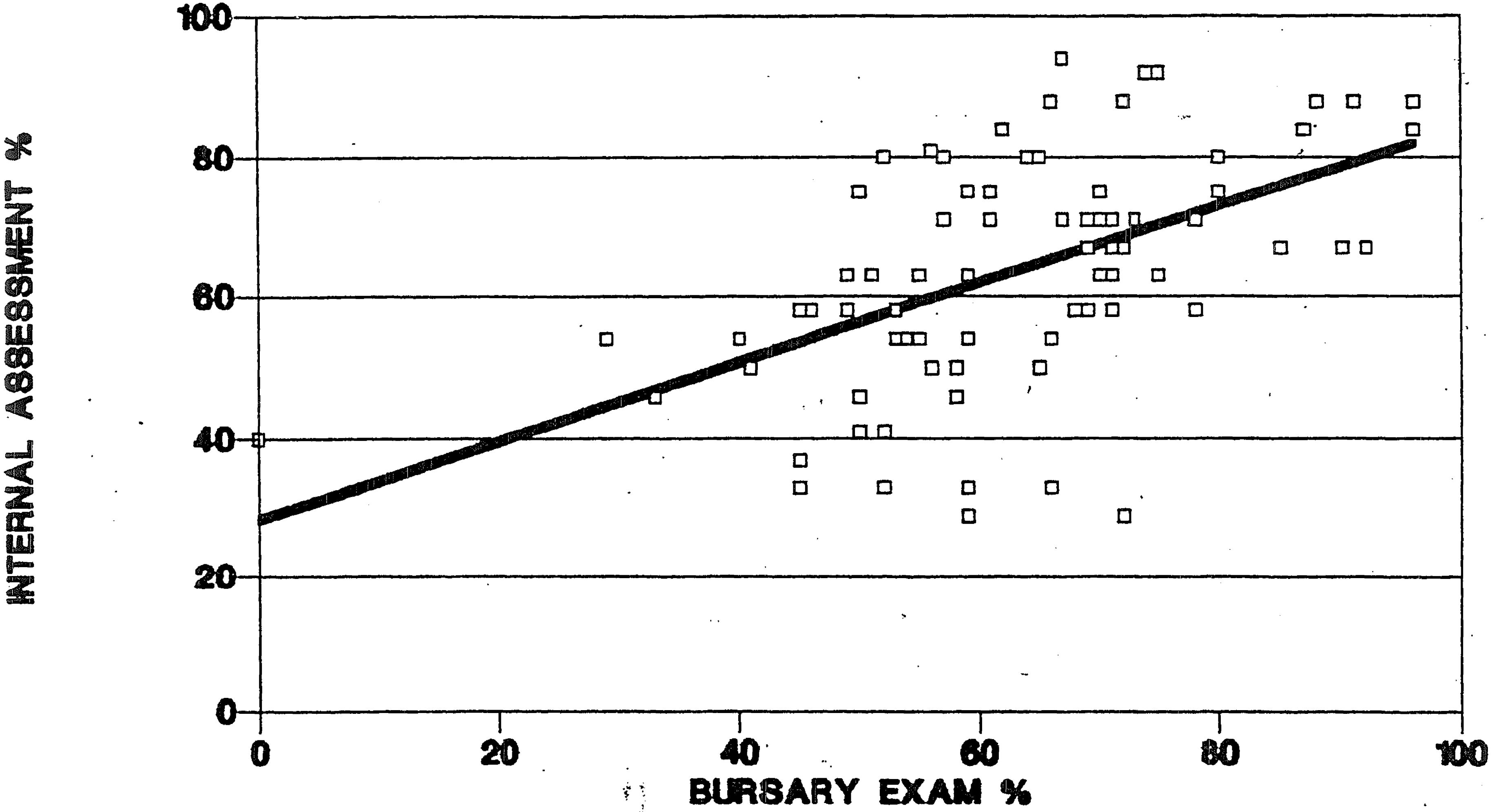
Pearson Correlation Coefficient

R = 0.28

Regression Line At

Y = 0.42X + 29.47

Graph 8 **Bursary Exam % versus Bursary Internal %**
(1988) (1988)



Pearson Correlation Coefficient

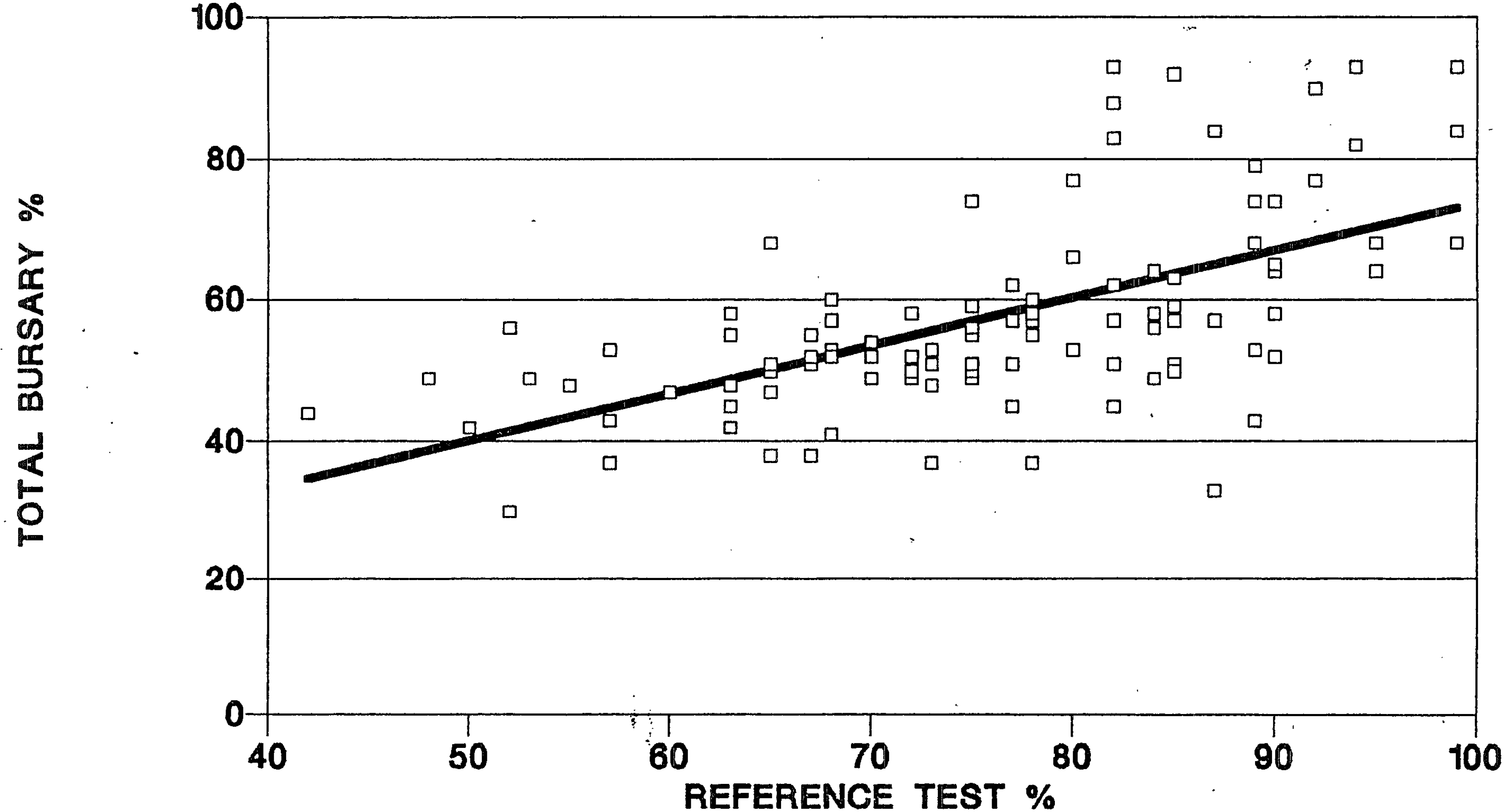
R = 0.54

Regression Line At

Y = 0.56X + 28.51

Reference Test % versus Total Bursary

(1987) (1989)



Pearson Correlation Coefficient

$R = 0.60$

Regression Line At

$Y = 1.13X + 6.28$