A STUDY OF CHEMISTRY 11 WITH SPECIAL REFERENCE TO THE EFFECTS OF THE SEMESTER SYSTEM AND SEX DIFFERENCE UPON ACHIEVEMENT

by

LAVINIA M. GREENWOOD
B.Sc., University of Victoria, 1969

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in the Faculty

of

Education

ACCEPTED
FACULTY OF GRADUATE STUDIES
[redacted]
DATE JUL 23 1979

We accept this thesis as conforming
to the required standard

[redacted]

UNIVERSITY OF VICTORIA
April 1979

ABSTRACT

Supervisor: Dr. L.D. Yore

A study of the Chemistry 11 course was undertaken with special reference to the effect of the semester system upon success in examinations. Three tests, entitled Forms A, B and C, were prepared on the Chemistry 11 course, and distributed throughout the province in 1972 and 1973. In all, 12,344 students were tested and graded by the Department of Education.

A sample of 2700 test papers was analyzed, using analysis of variance on the mean scores of students who had been enrolled in either the regular ten-month programmes or one of two different semester programmes. No significant difference was found between the three programmes using forms B and C. However, a significant difference between programmes was found for Form A, indicating that the mean scores of students from the first semester programmes in the 1972-73 school year were significantly lower than from the other two programmes.

The samples were also subdivided by sex, and a highly significant difference was found between boys and girls in their mean scores achieved in all three forms of the test. Boys scored higher than girls in all three replications.

Further study suggested that both these differences were related in part to the number of items in the test requiring mathematical manipulations. Boys and girls appeared to differ in their achievement on the test depending upon the way in which their chemistry and mathematics courses had been scheduled. Under some conditions girls scored higher than boys.

It was concluded that the semester system should be evaluated for its learning effectiveness as compared with the effectiveness of other systems. It was noted that when course schedules are prepared and course selections made, insufficient attention seems to be paid to the value of planning student's courses in a logical order of pre-requisites and co-requisites. Differences in the needs of the individual sexes should also be studied.

Examining Committee: [redacted]

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

This thesis has been an outcome of a project undertaken in conjunction with the Research Standards Division of the former Department (now Ministry) of Education. Between the preparation of the tests for the Division in 1971, and the completion of this thesis, considerable time has elapsed. Therefore I wish to acknowledge the contribution of many persons who have assisted me along the way.

Under the guidance of four different chairmen - namely, Dr. I. Walker, Dr. J.A. Downing, Dr. D.R. Daugs, and lastly, Dr. L.D. Yore, the committee membership has changed many times. There were those who made it possible for me to set out upon the road, and there have been other good samaritans without whom the journey would not have been completed.

I am indebted therefore at the start to the late Dr. L.J. Clarke who, while a member of my committee, opened the way, in 1970, for me to work with Dr. C.B. Conway and the Research and Standards Division. I am indebted also to the late Dr. Conway himself who subsequently made available to me such a wealth of data for this study, and answered so many questions: and my thanks are due to all the staff in his former department who have continued to supply me with pertinent information whenever I asked.

My special thanks must be preserved for two chairmen and advisors, Dr. J.A. Downing and Dr. L.D. Yore, without whom I would not have completed this thesis. By their unique assistance together with their personal encouragement, advice and patience , I have been enabled to fulfil the requirements of the university. In this regard I wish also to thank the Faculty of Graduate Studies in the persons of Dr. S.A. Jennings and Dr. J.M. Dewey for various extensions of time.

Many others have given me their advice and help. In preparing the tests I am indebted to the invaluable assistance of Dr. D.J. MacLaurin who spent so many painstaking hours checking and rechecking every item as the tests were written, and from whom I have learned so much. I also owe him thanks for staying longer than anyone on my committee from 1971-77 and so providing the only link with the beginnings.

Though never a member of my committee, Dr. R.R. Davidson of the Department of Mathematics has extended to me his invaluable advice and assistance. He helped me to unravel my sampling problems, develop my overall research design and who, through such generosity with his time, helped me to obtain a better grasp of the analysis of variance when undertaken with a variety of programmes.

Other committee members past and present deserve mention. Dr. W. Muir who in 1974, before joining my committee, assisted me in planning the original research design; and Dr. T.W. Dingle and Dr. R.L. Williams who have supported Dr. Yore in these latter months and thus assisted me in reaching the finishing post. My thanks go also to many others:

To Dr. J.D. Ayers for his advice on Test Theory. To Susan Rowntree without whose work in designing special programmes for my special needs, many errors would have gone undetected. To Mrs. Alison Armstrong of the Provincial Education Library who assisted me in searching her files and publications for information on the history of the semester system in British Columbia and Canada.

To Mr. A.L. Lamrock, formerly of the Greater Victoria School Board, who has so willingly supplied me with much valuable background information and with whom together with Dr. Glen Pope I have from time to time discussed the unpopular notion that girls may be different from boys.

To Miss Janet Blackmore who helped me to collate and code the data for the computer and to Mrs. Anne Field and Miss Diane Archer who key punched it all with such faultless accuracy.

And last but by no means least to my two typists - Mrs. Eleanor Lowther who typed and retyped these many pages again and again always ready with valuable advice on their presentation - and finally Ms Judy Pitcher who has been responsible for the finished production at very short notice.

DEDICATION

To my Husband -

whose support has never faltered.

CHAPTER I

INTRODUCTION

The Problem

In the last decade there has been a gradual introduction of the semester system into the secondary schools of the province of British Columbia. The arguments in favour of the practice are varied and are more related to student convenience than to a proven educational superiority of programme. It enables students to gather credits more easily but it is not known whether there is any effect upon the quality of learning that is taking place.

One way to evaluate the quality of the learning experience is to require students to write examinations. It used to be widely accepted that if one common examination is set for all schools, some valid conclusions could be drawn from the students' scores about the relative effectiveness of each school's programmes. In provincial examinations in British Columbia, there has been a tendency for students on the semester system to do less well than their contemporaries on the ten-month programme; and these differences have been greatest in the scientific subjects (B.C. Annual Reports, 1970-71).

While the problem which needs investigating is whether students learn an individual subject equally well on a five-month as on a ten-month course, it is only their achievement in a common examination which in practice is taken as evidence of their mastery of the course material. Such tests are simple to score and supply convenient data for statistical analysis. The larger question of what counts as having mastered a subject remains unasked by many people, because it can only be evaluated in the long term and in ways not amenable to statistical treatment.

For the purposes of research, comparisons between two schools' sets of internal results are of questionable validity where different teachers have set and graded their own tests. Accordingly, it was decided to prepare a test for distribution throughout the province. The subject matter was the course content of Chemistry 11 and the distribution was undertaken in cooperation with the Research and Standards Division of the Department of Education.

Procedure

In preparing the test, the writer was limited to a test length suitable for a one-hour classroom test. Since the whole course could not be adequately covered in so few items, three tests were randomly distributed throughout each class. This procedure has provided a means of replicating the experiment.

The Department of Education Research Division administered and graded the tests, and withdrew a sample of 3,600 papers for more detailed analysis. From this number, which comprised approximately one-third of the test population, 2,700 were identified as the data base from which to investigate the following questions:

1. Is there any significant difference between the mean test scores of students who were enrolled in the first or second semester programme or in a regular ten-month programme?

2. Is there any significant difference between the scores of boys and girls for the same programmes, and does any such difference change as the programmes change?

Limitations and Assumptions

In the development of this study, neither the final format and content of the tests nor the methods of distribution, sampling, and analysis were matters for decision by the writer. No preliminary trial distributions were made, so that there was no opportunity for amendments to weak items. A few errors were corrected before distribution in January 1973, but not all. A number of personal details about each student were recorded on the front page of each test including age, sex and related science and mathematics courses taken. Unfortunately, neither enrolment in Chemistry 12 nor repetition of Chemistry 11 could later be identified.

The method of sampling is described in Chapter III (Conway Letter, 1976, Appendix A). The sample was stratified with respect to the different programmes of study and undertaken systematically in order to obtain equal numbers in each stratum. It has been assumed that the sampling was sufficiently representative to cause no undue bias in an analysis of variance. Some data were discarded; only three out of the four programmes sampled were used in the analysis of variance. It was also assumed that the tests were distributed and written under the conditions requested by the Department of Education's Research and Standards Division in their procedural instructions (Appendix A).

In undertaking a wide survey of all schools in the province, other assumptions were numerous. For example, it is usually assumed for statistical purposes that all students have taken identical courses under equally competent and well qualified teachers, and equally good laboratory facilities. It has also been assumed here that previous education and exposure to related subject courses have not biased the students mean achievement scores. Such assumptions have been implicit in most educational research and are seldom questioned.

CHAPTER II

REVIEW OF THE LITERATURE

Two topics are reviewed here. The writer searched the limited available sources to try to discover reasons for the beginning of a semester system, and to survey attempts to evaluate its effectiveness. Both British Columbia and other sources in Canada were reviewed. The work of other researchers who have explored the teaching of school science to students of both sexes is also reported. An analysis of the conceptual demands of the Chemistry 11 course, known as Chem Study, together with a brief history of its development and introduction into British Columbia is contained in Appendix C.

The Semestered School Year

The semester system at secondary school level may be described as the practice of dividing the school year into two or more equal parts and presenting the student with a proportion of his required courses in each semester. In most schools in British Columbia, two five-month semesters have been introduced in which students take seven courses - four in one semester and three in the other. A few may take six or eight.

History

There appears to have been no formal beginning to the practice of semestering. The Chant Report of 1960 makes no mention of the idea. The first official reference to it in the B.C. Department of Education Annual Reports occurs in the 1969-70 edition under Field Services, in which it was simply noted that the practice was increasing. Nanaimo began a programme in 1968 in one school and was the first on Vancouver Island. Only one school in Ontario had tried it before 1969.

There seems to be no official public records for British Columbia of any formal decisions or plans. Tacit government approval may be inferred from an unpublished committee report (Department of Education, Canty, 1969) to the Minister in 1969 on school utilization. Here it was observed that "there is some merit in the development of the semester system of organization throughout the secondary schools, with an extension of the summer organization as an introductory step in this development" (p. 9). The intent was to suggest that there should be better year-round use of available facilities, with two semesters and a summer school, extending ultimately to a year-round school of three equal semesters. As well as the apparent economic advantage of making use of an expensive plant, it was hoped that more flexible transfer arrangements for students and entry into school more than once a year could be benefits of the new organization. The additional operating costs were not explored.

Neither was mention made in government circulars of more specific educational advantages. The initiative appears to have come from local districts. Nanaimo described their reasons for trying it in their report to their Board of Trustees in 1968. They had two sets of students they particularly wished to help - those who left grade 12 with one or two courses short of graduation requirements, and those who left school in the spring for commercial fishing or other work. A semester plan would attract students back to take just a few courses in one semester to complete graduation, and could offer credit for work done in a first semester to those who had left for work in the spring. Previously these students had been faced with repeating a complete year and so they seldom returned at all. It was also hoped that students who were dependent upon the school buses could get help from teachers during the school day in their study blocks.

In most schools, course organization began with two equal semesters with four courses in one semester and three in the other. Some schools had four 70- or 80-minute periods each day - others had five one-hour periods, so that the course in the first hour was repeated in the last hour of the day. In the one semester, the student had a number of free periods which could be committed to an extra course - perhaps one failed in the previous semester. Physical education was treated as one course, and course loads tended to be very uneven in the first years.

These obvious weaknesses, together with criticism of 70- and 80-minute periods, led to a 5 x 5 arrangement in some schools - five one-hour periods with five subject blocks. This arrangement gave the student the chance to take up to 20 courses in two years, while also spreading the heavier grade 11 and 12 courses - especially the sciences - over three semesters instead of two (Abbotsford, 1974). Students completing one semester would be given credit for grade 11 in that subject, but would receive no grade 12 credit unless the next two semesters were completed. Since the whole of the grade 11 course is not completed in one semester in this plan, such students who wrote this province-wide test of Chemistry 11 would be expected to be at some disadvantage in comparison with those from other schools. It has also been argued that spreading the total course of 11-12 sequence over a longer time is necessary because grade 12 content is more demanding.

Evaluation

Very little material seems to be available before 1974 on the topic of semestering. The studies which have been done have been undertaken by school districts for their own internal purposes and have not been published in the literature. In British Columbia the reports of School Districts 68 (Nanaimo, 1969) 61 (Victoria, 1970), 39 (Vancouver, 1971), and 34 (Abbotsford, 1974) were made available to the writer. The first three of these are in the form of opinion polls. Questionnaires were sent to teachers, students, principals, and parents; and from a wide variety of comments, a number of common opinions emerge.

1. The semester system does seem to have the approval of a high proportion of parents, teachers, and students. Students like the change of courses and teachers mid-year, and the chance to graduate in January. They claim that some courses are easier to master when in class daily, but others are clearly too condensed and the pace too great. The chance to repeat a failed course is appreciated,

but the opportunity is not as often made available in practice as in theory.

2. The greater pace makes good attendance essential. Some teachers report better attendance, others that students who get behind give up altogether.

3. There is an increased workload on teachers and administrators; but in spite of this, these groups are in favour of the system for its greater flexibility of programming for students and its success in attracting students back to finish their grade 12 requirements. Many respondents criticized the treating of Physical Education in this way, and the unbalanced workload on the students gave some schools problems with the use of the unscheduled periods.

Conflicting comments on whether the work and the homework are more or less arduous were clearly dependent upon the subjects in which the student was involved. The 1970-71 Annual Report of the Province referred to a growing concern over Mathematics and Physics under this system. There has been little study of the student's mastery of the subject. A writer of the Victoria study commented that "frankly no one knows what will be the effect of taking French 11 in the first semester in grade 11, and French 12 in the second semester of grade 12" (p. 3).

The Nanaimo study (1969) does report some comparative figures of passes and failures between 1968 and 1969. It was claimed that there were no significant differences in departmental examination results except that "because of repetition, more students were recommended." However for Chemistry 12, there were in fact noticeable differences. In 1968 (on the traditional ten-month arrangement), 26 students were recommended and 24 out of 28 passed the departmental exam. In 1969, 78 were recommended and only 10 out of 23 passed the exam. For Mathematics 12 and Physics the findings were similar, but at the grade 11 level, substantial increases were noted in the number of Mathematic and French passes awarded in the schools as a result of repetition of the courses. Such increases were seen as demonstrating marked improvement in achievement in these subjects and a justification for the semester system.

Later comments from both British Columbia and Ontario have been less enthusiastic. MacLennan (1974) complained about the need for good attendance and that sickness led to dropping out, but that students later could return. His school had used 80-minute periods which were found to be too long.

An extensive survey of semestering in Ontario was undertaken by Hill (1974). Only one school in the province was on the semester system before 1969, since which time the numbers had grown from 6 in 1969 to 52 in 1973. Hill wrote that he believed the rate of change would slacken as there was growing criticism and antagonism in those schools which had not changed "because of perceived educational disadvantages." He suggested that "only severe economic constraints would force these schools to adopt a semester system" and that it is most attractive to large urban schools and to separate schools, "possibly because of economies." The suspicion that semestering is a phenomenon benefiting only administrators and the budget comes through in a vigorous attack by Saunders (1974). He wrote:

> Semestering, in any of its variations, will permit staff reductions either by creating a situation in which students take fewer credits or by drastically reducing the number of instructional hours per credit. This fact alone would indicate that if semstering is analyzed on a 'benefit to whom' basis, that the principal benefactors are administrators at the expense of both teacher and student (p. 205).

In view of Nanaimo's stated goals, this seems to be an unfair view of administrative motives, although it is unlikely that the plan would have been implemented if it had been shown to be more expensive. Saunders also argues that the flexibility would require rehiring staff twice a year and that staff availability and scheduling severely limit this feature. Both writers strongly recommend a study of educational results before a "jumping on the bandwagon" by more schools.

In November 1974 Norman Gleadow undertook a more extensive evaluation for the Vancouver School Board. In addition to the findings already mentioned, Gleadon confirmed Saunders criticism by recording the number of schools which reported having modified the hours of instruction, and/or the course content, by either rewriting the courses or omitting whole sections of the course to fit into one semester, and courses requiring the understanding of concepts suffered. Hours of instruction per course varied from 60-180 hours, with six schools reporting between 90-95 hours. The lack of continuity was particularly deplored by teachers of Physical Education, Music, and Industrial Education. Librarians complained of lack of space to accommodate the larger numbers of students having non-instructional time to spend on study, with no place but the library to go, for Industrial Education the double ordering of supplies had increased the costs of this course.

Abbotsford attempted to solve some of these problems by more imaginative scheduling. Their report (Abbotsford,

1974) was an internal one by teachers seeking some solutions to complaints such as those expressed above. In particular they wished to reduce the effects of absenteeism, extended gaps between sequential courses, too rapid a pace in some subjects, and dislike of long periods. They sought to introduce a 5 x 5 plan mentioned earlier and spread the sequential 11-12 courses through three semesters. The report admits that the time for some courses will be reduced which bears out Saunders' complaints to some extent. If students do not proceed to a grade 12 course, one semester of grade 12 programme would receive no credits under their arrangements.

One study examining results has been undertaken in Victoria by J.R. Anderson in 1972. He studied the achievement of grade 11 students in relation to attendance and used students from schools on both the traditional and the semester system. His population was 1,402 grade 11 students from three schools in District 61, and his dependent variable was the Grade Point Average (GPA). From a sample of 180, he concluded that any difference in achievement of students on the semester and traditional systems was not significant. However, the results suggested that the effect of poor attendance on achievement for students in semestered schools was marked enough to require further investigation. By using the school Grade Point Average, no valid comparison can be made between schools at any subject level since school grades are arrived at only in relation to the general class standard, or perhaps that of several classes within one school; hence the purpose of periodic grade-wide examinations which have been traditionally set by the Department of Education to obtain some check on standards. The unpopularity of this practice among teachers may be partly due to the style of the examinations in the last 10 to 15 years, but there are some who consider their withdrawal a disservice to the students. Lief Stole of Edmonton, an Assistant Principal and former supervisor, has noted increases from 300 per cent in the number of 'A's'awarded in grade 12 matriculation subjects since the cessation of Departmental Examinations. This occurred in schools not previously noted for their higher academic standards and did not include the startling increase of 1,391.7 per cent in 'A's' in mathematics. Stolee (1978) claims that these increases are placing students from schools with more rigorous A grade standards at a disadvantage when applying for places in professional schools. A province-wide survey will allow the investigation of standard of achievement across schools.

Overall, the most constantly claimed advantage for the semester system was greater flexibility in course offerings. There seems to be some indications that quality may be being sacrificed for quantity.

Science in School and College and Differences Between the Sexes

i. Biological and Social Factors

In the emotional climate of today's feminist movements, it becomes difficult to suggest that there may be differences in the learning patterns of girls and boys. It is still more difficult to propose that there could be any value in investigating them and in making allowances in the running of schools without arousing an emotional reaction, unless it is in the area of elementary school and reading "readiness". Much has been both written and taught regarding the supposedly late development of boys and their resulting difficulties in reading and writing. Much of current learning theory is based upon a commonly held assumption that a male is physiologically one year younger than a female at birth. However, recent research has revealed that the difference in reading achievement which has been noted between boys and girls in Canada and the United States, is not universally observed. Downing and

Thomson (1974) have shown that in other countries, boys can do as well or better than girls. In Germany, boys are usually considered to learn to read faster than girls (Preston, 1962). Such observations make a biological explanation untenable and suggest a need to study cultural influences. Downing et al (1977) have found evidence of a relationship between sex role standards and an interest in and achievement in reading. They point to pertinent medical research on hermapheodites to show that a child's psycho-sexual identity is conditioned in early childhood.

A similar supposition, which persists in North America, regarding an innate inferiority of girls in mathematics and science, has also been questioned. An observer of Soviet education wrote that "there is not the slightest evidence that girls are any worse at scientific or mathematical subjects than boys" (*New Statesman*, 1956). Studies done on women's colleges in the United States record the observation that more women elect science courses in women's colleges than in co-educational institutions, where they fear to be outnumbered in class by men (Newcomer, 1959). At Vassar College for women, the percentage of women in sciences was higher a hundred years ago than it is today. This observation will come as a surprise to many who believe that we are more emancipated today than in the previous century with regard to the education of women. Harriet Martineau (1837), travelling in the United States in the 1830's, wrote

> We find it taken for granted that the girls are not to learn the dead languages and mathematics because they are not to exercise the professions where these attainments are wanted: and a little farther on we find it said that the chief reason for boys and young men studying these things is to improve the quality of their minds.

If women at Vassar were up against these attitudes a hundred years ago, is it still career expectations which keep girls and women out of science courses today when educational opportunities are now theoretically equally open to both sexes? Betty Friedan (1963) has shown in her book *The Feminist Mystique*, the social forces which have kept women out of certain careers; and Warren Farrell (1974) has spelled out the social conditioning processes on the outlook of boys in North American society. All societies naturally perpetuate themselves by example, and so influence the attitudes of their children. It is only the explanations which distinguish today's educationalists from their forefathers.

The problem is not confined to North America. From Australia, Keeves in 1973 wrote

> It seems unsound to propose physiological or psycho-analytical explanations for the difference in achievement and attitude in mathematics and science between the sexes when there are

> disparities between countries in the magnitude of the sex differences reported. The most likely explanation would seem to arise from the expectations which 'western' communities hold for the respective roles of men and women (p. 60).

He also believed that a perceived limitation in career choices influences the relative participation of the sexes in science activities, not only in universities, but also in the secondary schools (Keeves, 1973, p. 61). Opportunity and participation would appear to go hand in hand now as in 1830.

ii. Teacher Attitudes and Their Influence on Student Achievement

The importance of adult attitudes in the student's selection of career choices is obvious; but the influence of teachers as school counsellors, when they advise students on course and career choices, is perhaps not often questioned as to the possibility of sexual biases. Within the classroom teacher attitudes can reinforce or delay development. A belief that girls are not good at these subjects may quickly be conveyed to a class.

There are undoubtedly men today who still share Napoleon's view that "nothing is more detestable than a woman who reasons." Perhaps there is a connection here between a teacher's attitude and Crawley's (1975) suggestion that girls are, whether by nature or nurture, more respectful of authority. Whatever the cause, the obervation has implications for education, and some grounds for asking whether boys and girls may have different cognitive styles. Newbury (1966), writing on the teaching of chemistry, noted that while girls prefer demonstrations, boys like to do the experiments themselves.

Lodge (1969) asked how students rated the acquisition of certain science skills. He found that girls rated highest the skills of classifying and recording, while boys rated measuring and evaluating as most important.

As writers continue to speculate on the way children learn, the influence of elementary school experiences cannot be ignored. If girls on this continent seem to find it harder than boys to think through logical concepts in science and mathematics , is it because they found memorization a surer way of getting good grades in elementary school than asking questions and thinking things out? Crawley's observations are not trivial.

iii. Cognitive Styles In Secondary School Science

Field and Cropley (1969) made an extensive study of cognitive styles in relation go the learning of science subjects. They classified under the heading of *category width*, the concept of a student showing a *willingness to* *use broad inclusive categories, or narrow exclusive ones.* They found this variable correlated significantly with originality in which they found boys notably superior to girls. They also found boys to be superior in their development of *formal operations* as defined by Piaget (Inhelder & Piaget, 1958) and suggested that this may be a contributing factor to girls' inferiority in handling science information at school. High science scores for girls correlated surprisingly with the *narrow category width* - an opposite result from the boys in which high scores correlated with originality and flexibility. These writers go on to suggest that the stage of mental development of the pupil has an important bearing on performance, and that the sexes do develop differently. Shayer and Wylam (1976) investigated this function in two studies among British children between the ages of 9 and 14 years and later 14-16 years. They measured the proportion of children in their representative samples which showed different Piagetian levels of thinking. In summary they concluded that

> No increase in the proportion showing formal operational thinking was found beyond the age of 15. On the three tests used, a retrospective analysis of both surveys for sex differential showed no increase for girls after 14, while boys continued for a further year . . . the girls performance was substantially lower on the test

> on spatial relationships, and on the test on volume and density throughout the age range 9-16 (p. 62).

They go on to say that " . . . it must remain an open question as to whether there is further intellectual growth, as measured in Piagetian terms, at 17 or 18 or later years."

iv. Student Attitudes

Declining enrolments in science courses in both the United States and United Kingdom has been the spur to research into student attitudes. Course satisfaction and grading practices have therefore come in for questioning. Both Bridgham (1973) and Welch (1969) observed that science courses usually carry an overall lower grade average than other subjects so that students who select science courses must accept a kind of grade penalty in doing so. Bridgham suggests that if science grading practices were raised to compare with other subjects, 20 per cent more girls would enrol in chemistry and 80 per cent more in physics; but he suggests that the increase for boys would be minimal.

Bradley and Hutchins (1973) asked whether subject choice in secondary schools was related to a student's concept of science and scientists. They found that boys were much more favourably disposed towards science than girls, even at 13 years old, and felt that they were "doing well" at the subject at school. Girls tended to see science as something in which they had to learn too many facts, and science teachers were less popular with them than with boys. However, these same students did not themselves see science as being any more "difficult" for girls than for boys. It should be noted that these are 13-14 year olds from 17 schools in England and Wales, not senior secondary students.

These writers found no differences between the sexes in their attitudes towards scientists. However, the absent-minded professor has gone, along with other mythical pictures; and the typical scientist is viewed by both sexes as being an interesting, kindly, and a family man. There was no mention of anyone seeing a scientist as a woman, however, in spite of Mme. Curie's fame.

Gallagher (1969), on the other hand, found that although girls demonstrated a more favourable attitude than boys towards scientists, boys showed a more favourable attitude than girls on "myself as a scientist" questionnaire. He also found no differences between boys and girls towards science as a subject. These opposite results to the findings of Bradley and Hutchins may be partly explained by age differences, and partly by cultural ones. Gallagher was studying high school students in California and not English 13-year-olds.

v. Other Approaches

Other attempts to find reasons for sex differences in science achievement led to a study by Rothman, Welch and Walberg (1969). They attempted to relate student learning to the sexuality of the science teachers, but they were somewhat inconclusive as to the influence of these factors upon the achievement of boys and girls in their classes.

Summary

Falling enrolment in school science courses in western nations of the English speaking world has prompted research into attitudes towards science among school children. While achievement studies continue to show boys in the lead, relationships have been noted between achievement and attitudes, and between attitudes and cultural role expectations for each sex. The relative participation of the sexes in science subjects differs widely across the world, and seems to be related to career opportunities in the west.

In British Columbia, there is a question whether boys and girls are different in their reaction to school organization, to the presentation of subject matter and to the relative balance of related courses. How these questions are related to achievement in Chemistry 11 is the purpose of this paper. The introduction of a semester system has altered not only course choices, but the order in which those courses may be taken.

Readings in the semester system show that the programmes have been permitted to evolve for administrative reasons without clear educational justification. The main arguments used in its favour point to the desire to make it easier for students to acquire credits towards graduation. There is to date a general lack of scientifically controlled data on the effectiveness of the semester system. The quality of learning is questioned by teachers in Ontario, but in British Columbia these objections have seldom been raised until now. It should be noted that in the British Columbia school system, Chemistry 11 is an elective subject.

CHAPTER III

THE RESEARCH METHOD AND PROCEDURES

The Preparation of the Tests

The tests were prepared in the fall of 1971 under the direction of the Research and Standards Division of the B.C. Department of Education. The plan was for three forms of 30 items, and the maximum time allowance was to be one hour (Appendix A).

Content validity was the main aim of the writer in compiling the tests. This required full coverage of Chapters 1 to 8 and Chapter 18 of the textbook (Pimental, 1963) and equivalent parts of the laboratory manual (Malm, 1963), together with some additions to the course outlined by the Department of Education Curriculum Guide (1966). An attempt was made to design items which would test both mastery of the subject matter and the ability to make deductions from known and observed facts. The tests were therefore designed to include items of a wide range of difficulty, from some requiring simple memory or acquaintance with the topic to those which expected an ability to apply a principle to an unfamiliar situation.

Items were chosen so as to cover a similar topic or concept on two of the three forms and were placed in different positions within the two tests. By this device it was expected that a topic frequently omitted in one form, because of its position near the end, would be attempted in the other test. Approximately half of the test items required calculating a numerical answer.

An attempt was made to balance the three forms in their overall difficulty, and in the proportion of mathematical, true-false, and one answer type items which each contained. The majority, about two-thirds, were in multiple choice style. A number of errors were present in all three forms on distribution, which are noted in Appendix A. Students were also asked to provide information on the front page of the test forms, concerning not only their personal data, but also of their past or present enrolment in Mathematics 10, 11, 12, Biology 11, Physics 11, and Science 10. Unfortunately, the question of enrolment in Chemistry 12 or repetition of Chemistry 11 was not included (Appendix A).

The Research and Standards Division of the Department of Education undertook the task of printing the tests and distributing them throughout the Province on two occasions.

In June 1972 they were administered to all available students who were taking or had taken Chemistry 11 in the 1971-72 school year. In January 1973 they were administered again (after the correction of some typing errors noted in Appendix A) to the new set of grade 11 students who had enrolled in Chemistry 11 in the first semester of the 1972-73 school year. The distribution of papers within each classroom was to be undertaken in accordance with specific directions which began by stating the following purposes of the survey:

> These tests are general achievement tests at the Chemistry 11 level and are designed to measure the extent of a student's mastery of content and applications of Chemistry 11 material. Losses and gains in semestered and non-semestered schools will be compared. Each student will receive one of the three forms which are to be treated as one for administrative purposes. After the norms have been calculated the tests will be available to teachers of chemistry students.

There followed instructions for distributing the test papers:

> The examiner probably will have noticed that three forms (A, B, C) of the Chemistry 11 tests are alternated in the package. This will ensure random sampling throughout the school and throughout the province. Tests are to be distributed as they are received; *the distribution should be made as if there were only one form.*

The arrangement which is referred to was an interweaving of equal numbers of test forms A, B and C " . . . so that random samples of students would be obtained in each class, and the equivalence of the three forms could be determined" (Conway, 1973, Appendix A). The working time of the test was 60 minutes. Scratch paper was provided to the student at this researcher's request.

Population

The three test forms were randomly distributed in the above way to a total population of 12,344 Chemistry 11 students. This number does not include all students who had initially registered in the course since some students had left school prior to the examination date. In 1972, 2530 students who had been enrolled in September 1971 in the first semester of that school year were omitted from this analysis because it had already been demonstrated that this group's achievement was significantly below the rest of the population (B.C. Department of Education, Conway, 1973). The resulting population from which a sample was analyzed numbered 9814 students.

Table I displays the number of students in each of the remaining three groups, and the percentage of the smaller population that is represented by each group. It will be noted from the Table that in 1972, less than 40 per cent of the province's senior secondary programmes offering

TABLE I

Number of Students in the Population of Groups R S and F,
And the Ratio of Each Group to the Total

Programme	R	S	F	Total R S F
Number enrolled	3700	2668	3446	9814
Number as % of total enrolled	37.7%	27.2%	35.1%	100%
Number as % of total population tested	30.0%	21.6%	27.9%	79.5%

R 3700 students enrolled in Chemistry 11 in the regular ten-month course (September 1971 - June 1972).

S 2668 students enrolled in Chemistry 11 in the second semester from January to June 1972.

F 3446 students enrolled in Chemistry 11 in the first semester from September 1972 to January 1973.

chemistry courses remained on a ten-month programme (R group). The other 62 per cent were divided unequally between the first and second semester (F group and S group).

Preliminary Treatment of Subjects and Data

The Research and Standards Division collected and scored the entire population of tested students. Their purpose was to establish provincial norms, and also to make comparisons between the semestered groups and the ten-month programme and between groups with differing mathematical preparation. For this the Division used mean scores from the total population. They also selected a sample (in the manner described in the next section), after grading the papers for the purpose of making a detailed item analysis and exploring variations in the difficulty evidenced by certain topics. These 3600 papers were then made available to the writer.

Upon receipt of the papers, the students' test results were coded for data processing onto computer cards. Each answer was recorded and the test scores computed using the programme TESTAN (Appendix A), which also provided test statistics in the output. Information recorded by each student on the front page of the tests was also recorded on the computer cards to provide data for cross-classification. This double scoring enabled the writer to correct any errors that occurred in the scoring process by comparing the scores arrived at by each researcher.

Sampling Procedures

The selection of the sample was not undertaken by a simple random method, but was gathered in a stratified and systematic manner. Exactly 300 papers were withdrawn from each of the three forms for each of the three programmes, making a total of 2700 papers. They were withdrawn systematically from piles of graded papers which were stacked alphabetically by school, and any excess over 300 per form and programme were randomly discarded. Papers with zero scores from students who clearly had not attempted the paper were also discarded. Thus the sample was stratified with respect to the three forms and to the variable of main interest - the population groups described earlier.

This sampling was done only once. Thereafter a cross-classification technique was used to control and analyze those variables which were suspected of having the greatest influence upon the dependent variable. Categories which were recorded included the age and sex of the student, the mathematical courses taken or being taken concurrently, physics or biology courses, as well as school, school district, grade level, and other identificational data. Tables IX and X in Appendix B indicate that this method of sampling yielded a sample means which well represents the original population means, in one case, within one half of one per cent.

Sample Proportions

By selecting exactly 300 papers from each group in each of the three forms, unequal proportions of the population in each group were obtained. Of the students taking Chemistry 11 in the 1971-72 school year, approximately 40 per cent were enrolled in the ten-month courses, the remaining 60 per cent came from a semestered organization divided more or less equally between the two halves. Thus the proportion of students sampled from the regular ten-month programme was lower than from the two semestered groups. Table II shows that these proportions range from 23.0 per cent to 34.0 per cent. Other classifications, however, were obtained in the same proportions as had occurred in the total populations. Sex proportions turned out to be approximately 3:2, Boys:Girls. Table III shows the number and proportions of boys and girls in each group.

TABLE II

Sample Percentage: Total Numbers (N) of Students Writing Tests in Each Programme and the Sample Ratios Taken (f_x)

Prog.		Form A	Form B	Form C	Total/Programme
R	N_1	1259	1235	1206	3700
	f_1	23.8%	24.2%	24.9%	24.3%
S	N_2	891	880	897	2668
	f_2	33.7%	34.0%	33.4%	33.7%
F	N_4	1168	1148	1130	3446
	f_4	25.7%	26.1%	26.5%	26.1%
Total per form	N_T	3318	3263	3233	9814
	f_T	27.1%	27.6%	27.8%	27.5%

Sample N = 300

$f_x = 300/N_x$ for each Form

$f_T = 900/N_T$ for each Form

TABLE III

Number and Percentages of Boys and Girls in Each Sample

Programme		Form A		Form B		Form C		Totals
		Boys	Girls	Boys	Girls	Boys	Girls	
R	N_s	176	124	180	119	181	119	899
	%	58.7	41.3	60.0	39.7	60.3	39.7	
S	N_s	164	136	174	126	163	137	900
	%	54.7	45.3	58.0	42.0	54.3	45.7	
F	N_s	193	107	168	129	168	131	896
	%	64.3	35.7	56.0	43.0	56.0	43.7	
Unknown		0		4		1		5
Totals		533	367	522	374	512	387	2700
Form %		59.2	40.8	58.2	41.7	56.9	43.0	

Treatment

Test forms. In this study, the only element of experimental control as defined earlier was exercised in the random distribution of the three forms of the tests in approximately equal numbers in each classroom (Conway, Appendix A). It was the intention of the Research and Standards Division to pool the results of the three tests if their test statistics showed them to be sufficiently "equivalent". In the opinion of the writer, the variances were not sufficiently homogeneous, so the three forms were treated as non-equivalent, and the three forms utilized as three replications of the one survey. The forms were labelled A, B, and C (Conway Letter, Appendix A).

Programmes

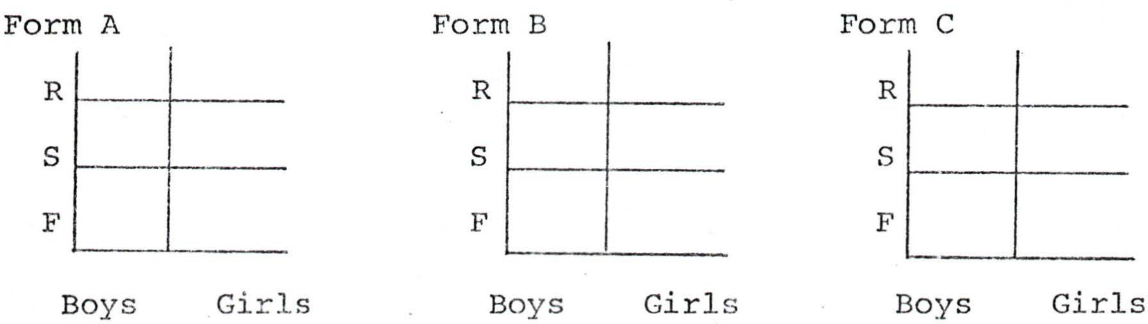
The first main variable of interest, which determined the choice of strata for sampling purposes, was the type of school programme in which the students took their Chemistry 11 course. The programme was either ten-months or a semstered school year, and students were enrolled in one of three programmes defined as follows: (a) Programme R: a ten-month programme for September 1971 to June 1972, writing the test in June 1972 (Regular Programme); (b) Programme S: a second semester programme from

February to June 1972, writing the test in June 1972; (c) Programme F: a first semester programme from September 1972 to January 1973, writing the test in January 1973. A *semester* is a period, usually five months long, in which a student takes only half the number of courses prescribed for credit for one school year.

The second main variable was the sex of the student. This variable was obtained by classifying the data, after scoring and sampling, from information provided by the student on the front of each test paper. Five students of Form B and one of Form C did not identify their sex, and the papers had to be discarded before undertaking the two-way analysis of variance.

Design

A two-way analysis of variance was performed on the mean scores of each group on each form of the test separately. The analyses took the form shown below:



Hypotheses

The following null hypotheses were tested:

H_1: there is no significant difference between the mean scores in programmes R, S, and F;

H_2: there is no significant difference between the mean scores of Boys and Girls;

H_3: there is no significant interaction between the sexes and the programmes in mean scores of each of the six subgroups.

Each of these hypotheses was tested separately for Forms A, B, and C. The analysis was undertaken using the computer programme ANOVA 22 which computes a two-way analysis for uneven cell frequencies. After reviewing the power function at $\alpha = .01$ and $\alpha = .05$, the level of significance was set at $\alpha = .05$ (Table XI Appendix B).

CHAPTER IV

PRESENTATION AND INTERPRETATION OF THE DATA

Test Statistics

Table IV records the mean scores, standard deviations, and number of students in each of 18 subgroups when the data are categorized by semester, sex and test forms. Figure 1 illustrates the mean scores in relation to the three programmes and shows the means with sexes pooled and sexes separated. It may be observed that the means are everywhere lower for girls than for boys and so are the standard deviations, variances and reliability coefficients. The only exceptions are the two reliability coefficients for Forms B and C in the second semester (Table IV, page 41, and Tables VI and VII Appendix B).

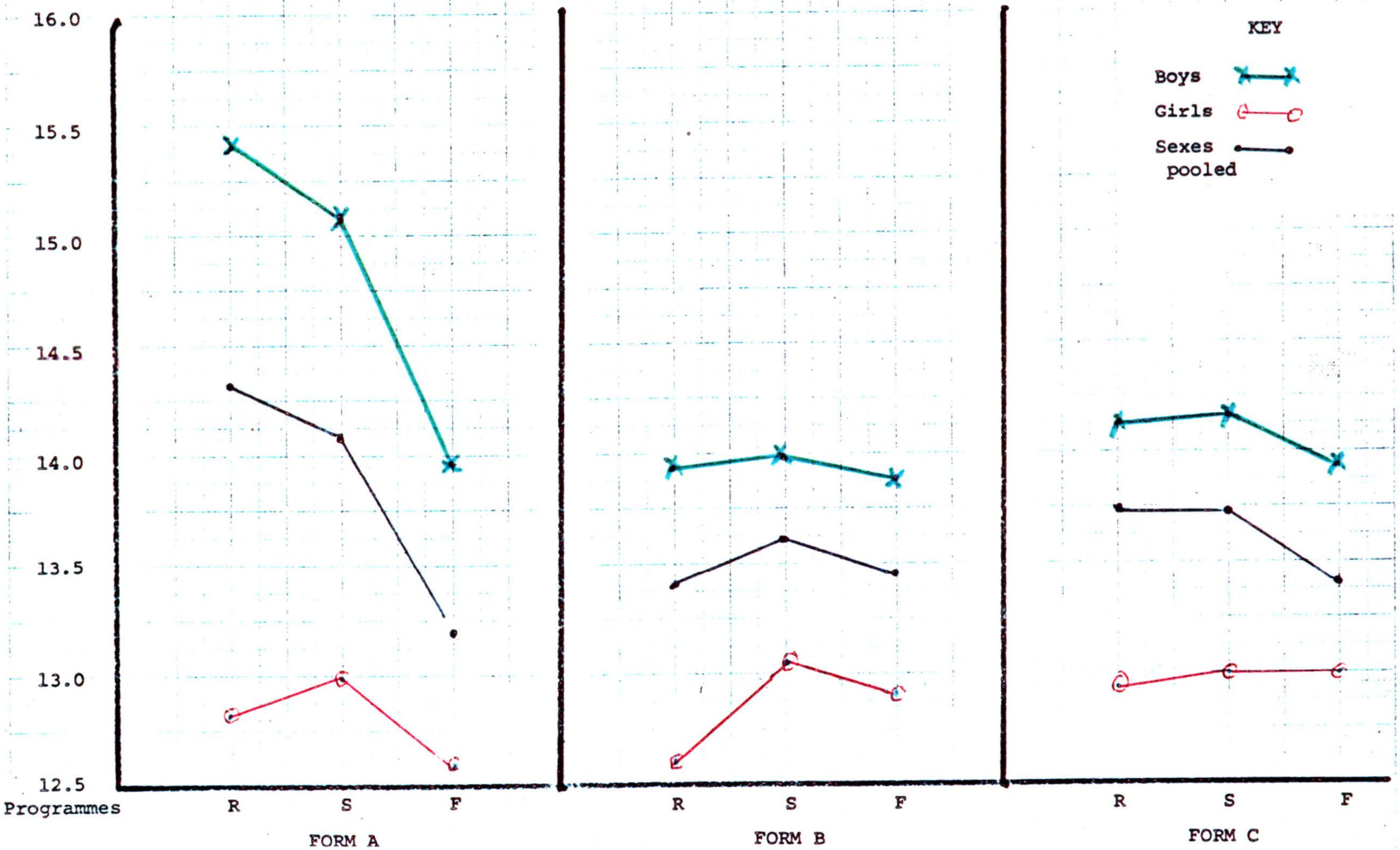
TABLE IV

Mean Scores (m) and Standard Deviations (sd)
and Enrolment (n) in Each Subgroup

			Boys	Girls
Form A	R	m	15.38	12.81
		sd	5.71	4.38
		n	176	127
	S	m	15.09	12.98
		sd	5.44	4.89
		n	164	136
	F	m	13.56	12.58
		sd	5.31	5.10
		n	193	107
Form B	R	m	13.94	12.61
		sd	4.51	4.19
		n	180	119
	S	m	13.97	13.07
		sd	4.19	4.19
		n	174	126
	F	m	13.86	12.89
		sd	4.56	3.92
		n	168	129
Form C	R	m	14.15	12.92
		sd	5.29	4.43
		n	181	119
	S	m	14.17	13.00
		sd	5.70	4.98
		n	163	137
	F	m	13.94	13.00
		sd	4.71	4.45
		n	168	131

FIGURE 1

Mean Scores Showing Sex Differences In Each Programme In Chemistry 11



Results of the Analysis of Variance

The two-way Analysis of Variance, using the programme ANOVA 22, yielded the following results

TABLE V

Analyses of Variance[1]

	Source	*df*	MS	*F*	*P*
FORM A	Sex	1	786.90	28.96*	0.00002
	Programme	2	134.26	4.94*	0.007
	Interaction	2	47.61	1.73	0.18
	Error	894	27.17		
FORM B	Sex	1	246.69	13.40*	0.0003
	Programme	2	3.21	0.17	0.84
	Interaction	2	3.88	0.21	0.81
	Error	892	18.42		
FORM C	Sex	1	269.82	10.81*	0.001
	Programme	2	1.22	0.05	0.95
	Interaction	2	1.65	0.07	0.94
	Error	888	24.95		

* Significant at the .05 level.

1 Based on Winer (1962, pp. 224-227, 291-297).

Interpretation of the Analysis

H_1: Programmes. For Forms B and C, there was no significant difference in the mean scores of students in the three programmes, R,S and F. In these two replications we cannot reject the hypothesis. Form A shows different results. The mean scores between the programmes showed a significant difference at less than 1 per cent. Therefore, for Form A, the hypothesis may be rejected (Table V).

The graph illustrates a tendency for programme F - the first semester tested in January, to show a lower mean score especially in Form A.

H_2: Sex Differences. Table V shows a significant difference between the mean scores of boys and girls on all programmes. This was true in all three replications. The differences are illustrated in Figure 1 and show the boys always in the lead. Therefore, Hypothesis Two is rejected.

The graph also shows that there was a tendency for girls to do best in the second semester in all forms. Boys seem to do equally well on a ten-month or second semester situation, but less well in a first semester. The significant difference between the programmes in Form A is located among boys, not girls.

H_3: *Interaction Effect.* From the results shown in Table V, the hypothesis that there is no interaction effect between the programme and the sex of the student cannot be rejected. Mean scores illustrated in the graph, however, suggest the presence of some interaction. They indicate that the significant difference between programmes for Form A is due to the boys' scores and not the girls' scores. This degree of interaction invites further investigations by methods of analysis which are designed to uncover underlying causes and explore several of the potentially interacting factors in a more specific way.

CHAPTER V

DISCUSSION AND SUGGESTIONS FOR FURTHER RESEARCH

Overview

In order to assess the relative quality of the learning environment in Chemistry 11 courses where programmes have been reorganized under a semester system, a set of three similar one-hour tests were randomly administered to the Chemistry 11 population throughout the province of British Columbia in June 1972 and January 1973. A sample of these test results was selected and analyzed by a series of two-way analyses of variance. Programme and Sex were the main effects for the analyses. The following hypotheses were tested: (1) that there were no significant differences in the achievement of students on the basis of sex, (2) that there were no significant differences in the achievement of students on the semester or non-semestered programmes, and (3) there were no significant interaction between the sex and programme factors.

The three forms of the tests - A, B, and C - though carefully matched in length and difficulty, were not identical in subject matter. All topics included were part of the Chemistry 11 course, and the majority were presented in two out of the three tests. The results were not pooled for analysis purposes because the difference in the variances of the three forms suggested that they should not be regarded as equivalent forms. They were therefore treated as three parallel studies.

The two-way analysis of variance yielded the following results: (1) There was a significant difference between the mean scores of boys and girls on all three test forms, the boys showing the higher scores in all three; (2) There was no significant difference between the mean scores of students in Chemistry 11 who took the course on a ten-month or on a semester basis for test forms B or C. For Form A, however, there was a significant difference in favour of the ten-month and second semester over the first semester scores; (3) There was no significant interaction between programme and sex indicated by ANOVA 22.

Discussion

The results from the analysis of variance show a difference between the three forms, but this statistical significance cannot by itself lead to any conclusions about the reasons for this difference - other information must be taken into account. Examination of Table V (page 43) and Figure 1 (page 42) reveals that the significant programme effect in Form A was due to the boys' scores and not to those of the girls. Also Figure 1 shows a tendency for girls to do best in the second semester, and boys to do least well in the first. These findings invite further investigation into possible masked interaction effects, as well as indicating points for discussion relating to the main effects.

Sex Differences

In Forms A, B and C boys performed significantly better ($p \leqslant 0.05$) than girls on those skills measured by the evaluation instruments used in this study. These instruments showed a lower reliability for girls than for boys. Further study of sex differences in the secondary high schools would seem to be indicated. The findings of other writers do not confirm the belief that girls are, in all cultures, inferior to boys in their mastery of scientific and mathematical subjects (New Statesman, 1956). The reasons must be sought in other areas such as the attitudes of students towards scientific careers, attitudes of teachers towards students in science classes, and in the general expectations which surround the learning environment.

Strong social pressures connected with role expectations for the sexes have been pointed out by Frieden (1963) and Farrell (1974). It has also been noted by Newcomer (1959) that the participation of women in science education differed with the social environment, being higher in women's colleges in the United States than in co-ed institutions. Thus the clear differences between sexes in their participation in school educational activities and in career choices may not reflect their relative innate abilities. Differences in achievement in science and mathematics, noted by Keeves (1973) in his international study in ten countries, showed widely differing results between nationalities. In analyzing the complex network of forces associated with cultural norms, expectations and sex roles, he suggests that insufficient attention has been given by researchers to the possibility of inequalities between boys and girls in the educational opportunities made available to them. Teacher attitudes both inside the classroom in their teaching styles and expectations, and outside in the planning of the time tables, may be a critical factor in influencing students' choices of courses.

Another potential factor that has been reported is the differences in cognitive styles between boys and girls (Field and Cropley, 1969; Lodge, 1969). Their findings suggest that teaching and testing methods may be more suitable for one sex than another. Newbury (1965) has noted that in laboratory teaching, girls prefer demonstrations while boys prefer to do the experiments themselves. Current testing methods strongly favour memorization, but in this study, the three tests contained a considerable number of items requiring applications of a principle to an unknown situation. It has been noted that the tests' reliability indices were lower for girls on all three forms suggesting that in some way the test instruments were more suitable for boys. Some might wish to argue that the emphasis on the application level of questioning that was used in this study's tests could account in part for the consistent sex differences in the achievement scores. If we can equate the application level of questioning with problem solving and the Piagetian stage of thinking known as formal operations, then these findings could be said to confirm more recent investigations of Shayer and Wylam (1976). After studying student achievement in problem solving, they concluded that a girl's development in operational thinking ceases or slows down temporarily around 14 years, whilst in the boys it continues to develop throughout their fifteenth and sixteenth year. Although their study leaves open the question of renewed development in later years, it does suggest a rationale for applying different teaching strategies to the secondary school age group. Some further indications of this possibility are present in the results of the analysis of the school programmes, to which programmes boys and girls appear to react differently.

Differences Between Programmes

For Form A there was a significant difference between programmes due to the boys scores. Figure 1 which illustrates the results in Table V, shows that the ten-month year is favored for the boys, but is only second best for girls. The results when the sexes are pooled still favours the ten-month programme over fall or spring semesters.

No such significant differences were found for Forms B or C. This does not imply that Form A results should be ignored however, but it does suggest that any inference which might be drawn from Form A's pooled results should be examined more critically. The inference that Chemistry students will achieve better results if the course is taught on a regular ten-month programme can be seen on closer examination to be only a justifiable conclusion for boys. Differences between the results of the three Forms raises more interesting questions than the simple one of which is best - semesters or ten-month programmes. Form A's significant trends favouring the ten-month second semester programmes over the first semesters can be detected to a non-significant extent in the other forms, supporting the need for further more analytical investigation into contributing factors. Although the results were clearly test dependent the overall achievement levels and trends among the programmes were comparable. The tests were not rated equivalent and no claim was made to achieve an item-by-item equivalence, but only a general standard of overall difficulty, range of topic and type of item included on each test. A closer look at the item statistics and the composition of Form A was taken in order to search for some underlying relationship between the significant programme effect and this test's particular pattern of results.

How was Form A different? It was the longest test in terms of the number of test papers which were recorded as unfinished. By definition, a test paper is unfinished if the last item or items are not apparently attempted. The substantial difference between the number of unfinished tests for the three forms (Table VIII) is directly related to this definition. Form A had two of the hardest questions as the last two in the paper. Forms B and C had items of similar difficulty placed in other positions on the test. By this definition the fewest unfinished tests occurred in the ten-month programmes and the greatest number occurred in first semester. Girls proportionally left more tests unfinished than boys.

Another way in which Form A differed slightly from the other two was in the number of items requiring mathematical calculation. Form A had 17 items, Form B, 14 and Form C, 15 items out of 30 in this category. These slight differences parallel the differences in the number of unfinished papers, suggestive of the importance of the time element involved in calculation and problem solving.

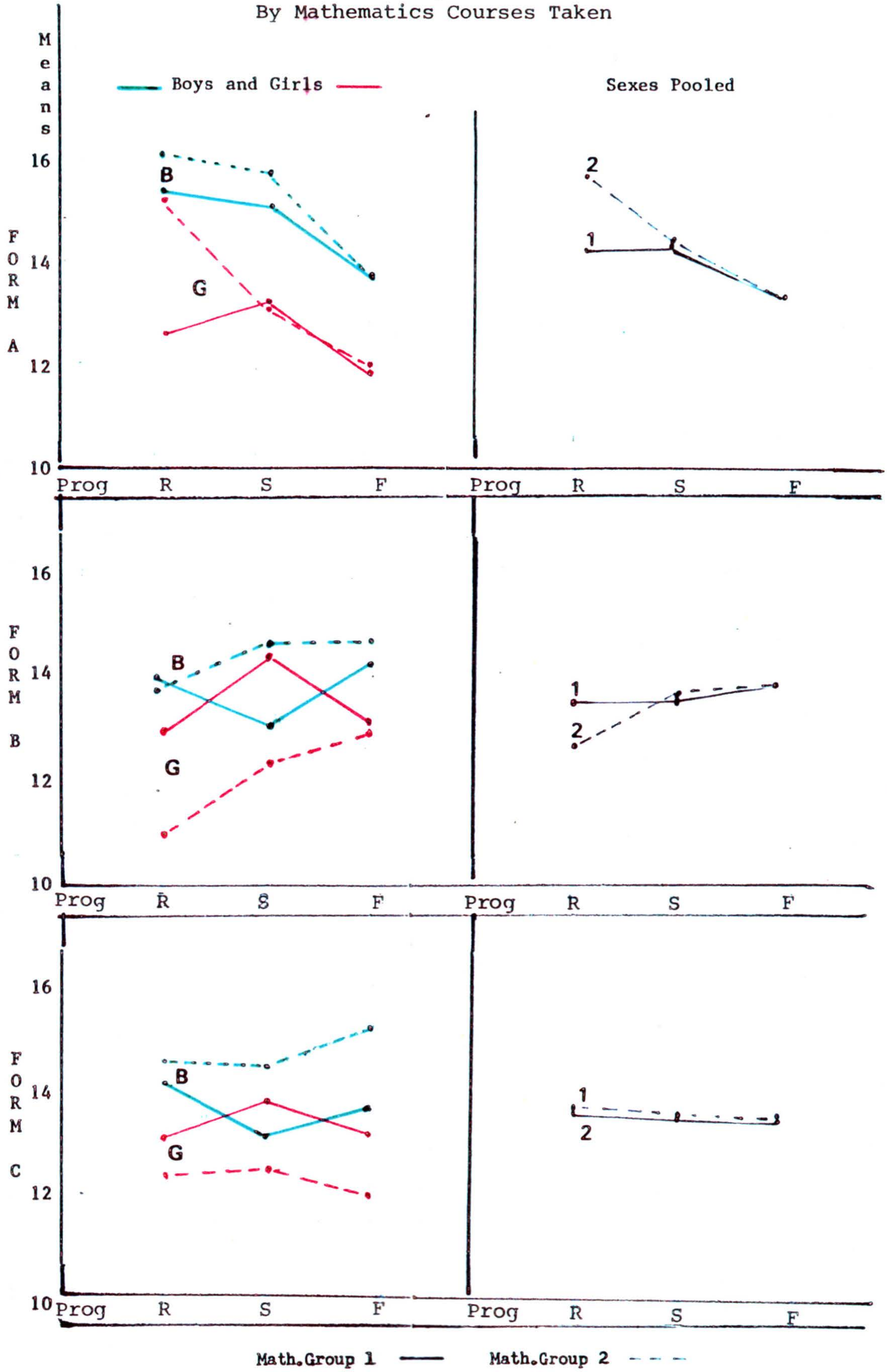
In 1973, Conway used difficulty and discrimination indices to uncover items of greater difficulty and showed that some of the more difficult topics were those in which mathematical manipulations were required (British Columbia Department of Education, 1973). In a second study he used the data from the entire 1973 population to assess the effect of the scheduling of Mathematics 11 courses relative to Chemistry 11 under the semester system. Conway divided the population into three groups - those who had taken Mathematics 11 before, during or after Chemistry 11. He concluded that although "it can readily be seen that those who previously [have] taken Math 11 usually have higher chemistry scores, a positive conclusion that Math 11 should be a prerequisite is not justified." He argues that the results are usually due to self-selection of pupils in terms of general ability.

The writer questions this argument since a student's choices, when they register for new courses, are as often influenced by what is available as by personal interests. They may also be guided by the advice of a counsellor, but there seems to be little evidence that chemistry students have been advised of the appropriateness of doing Mathematics 11 before or at least concurrently with Chemistry 11.

The writer's analysis of the sample means did not support the view that reorganizing the school year into short semesters makes no difference to student achievement. By breaking all the large groups down into the smaller subgroups by sex and mathematical preparation, some interesting trends and differences appeared which were obscured by the averaging process. A visual display of some results from a preliminary investigation into this aspect of the problem is shown in Figure 2. The points on the graph are taken from the mean scores recorded in Table IX (Appendix B). In Figure 2, in group 1 students who were taking Mathematics 11 concurrently with Chemistry 11

Figure 2

Mean Scores for Boys and Girls Divided
By Mathematics Courses Taken



and in group 2 were those who had completed Mathematics 11 before beginning the chemistry course. It would appear that boys do consistently better if they take Mathematics 11 in advance of Chemistry 11, but that girls have a greater measure of success if they take Mathematics 11 concurrently with Chemistry 11. The differences average out when the sexes are pooled, completely obscuring the importance of co-requisite mathematics.

The observation that some of the difficult topics were those in which mathematical manipulation were required would suggest a need to coordinate the teaching of science and mathematics courses when planning school curricula.

Lockwood (1957) investigated the general belief that high school students often avoided chemistry and physics because of the level of difficulty of the mathematical content. He concluded that a first course in algebra was sufficient for the needs of chemistry courses. Denny (1971) analyzed chemistry tests which were copyrighted in the 1960's, and listed ten mathematical skills that were needed to solve all problems that would be encountered therein. A review of B.C.'s mathematics courses in light of these ten needed skills indicates that five of them were not taught until Mathematics 11. The traditional year programme was planned so that Chemistry 11 and Mathematics 11 would be taken concurrently but under a semester schedule a student may enrol in Chemistry 11 and Mathematics 11 in different semesters. It was ascertained from Conway's records that the proportion of students in Programmes R and S who had not taken Mathematics 11 before completing the Chemistry 11 course was only 4 per cent, but in Programme F this proportion rose to 29 per cent. These students would have taken no algebra course since grade 9, since at that time the grade 10 course consisted solely of geometry. Some evidence that this weakness in mathematics grounding could affect chemistry scores may be seen by comparing Figures 1 and 2. In Figure 2 the latter mathematical group has been omitted, and the mean scores in Programme F are no longer perceptibly lower than in the other programmes, suggesting that this group may be one cause of the lower performance shown in Figure 1 and Table V. The large proportion of 29 per cent suggests a lack of interest in any joint planning of mathematics and science courses at several levels of the system.

Many writers comment on the lack of co-operation between mathematics and science teachers. Both Crawford (1971) and Rising (1967), in writing of this need, call for sustained and serious consultation between them.

Students do not necessarily recognize a familiar technique in a new guise; there are frequently differences in notation, terminology, and procedures between the three sciences due to traditions in each discipline. There is no mention made in the provincial chemistry course outline of the mathematical skills required for success in the course - an omission which illustrates the point. The disturbing feature of these observations is the prevalence of organizational practices which seem to take so little account of the students' learning needs, in spite of a large body of experience and research.

Interaction

There was no significant interaction effect found between programme and sex from the results of ANOVA 22.

Examination of Figures 1 and 2 indicate the potential for some interaction of programme and sex. Figure 1 illustrates that the significant programme effect for Form A was due to the boys' scores, and Figure 2 indicates that girls could gain higher mean scores than boys under certain programme conditions, namely, if they took their chemistry in the second semester and were studying their Mathematics 11 concurrently.

The above observations raise the question as to whether the selected method of analysis was the most appropriate in view of the manner in which the data was collected. The design was one which has been described as "quasi experimental" (Glass & Stanley, 1970, p. 501) because there was no random assignment of experimental units (students) to treatment categories. Instead the control of significant variables was undertaken by classification of the data after the tests were written and graded. Thus, the choice of an analysis technique which is intended for use with the true experimental design may not have been the most appropriate mathematical method, since it may lead to incorrect interpretation. Overall and Klett (1972), in describing three methods of calculating an analysis of variance, observed that

> powerful suppressor effects are not uncommon between classification variables and their interactions . . . . In dealing with highly interdependent classification variables it may not be useful to test the significance of the independent effect of each factor adjusted for all others. Method 3 has a pronounced advantage in minimizing the possibility that significant effects will cancel one another out (p. 454).

Method 3 referred to here is known as the *step down analysis* or the *forward elimination of effects* and can be used to show the total effect of a factor unadjusted for other factors. It is clear from the data presented, and Figures 1 and 2, that there are considerable cancellation effects presented in this data, both between programme and sex factors and between sex factors and the scheduling of mathematics courses. A three-way analysis using the step down approach could reveal the fuller extent of these interactions.

Another approach could be to eliminate one set of interactions by undertaking an analysis on each sex separately. By this method the differences in the variability becomes reflected in the separate F ratio for each sex when testing for significance of the programme effect. It follows that the chances of obtaining a significant F ratio will be increased with this data, because the larger population (boys) is also the group with the large variances (Glass & Stanley, 1970, p. 372; Overall & Klett, 1972, pp. 455-458).

In summary, despite the limitations of the programme which was used, it seems clear that further investigation which takes into account a greater number of variables, in particular the scheduling of related courses such as mathematics, might more clearly define what lies behind the programme factor.

What Further Research Is Indicated?

The following areas have been identified as requiring further research. Each could be studied individually, but more meaningfully if they are cross-classified with the sex factor.

1. Item statistics could be studied in detail to examine whether difficulty and discrimination indices are sex-dependent or test-dependent in terms of the position of the item on the test. They could also be evaluated in terms of the degree of mathematical preparation a student had received before taking the chemistry course or the chemistry test. The programme of studies in which a student was enrolled - ten-month or semester system - should be another factor upon which to compare the difficulty and discriminating power of a particular item. The dependence of test writers upon such statistics in designing tests makes this type of research imperative.

2. Further examination of the semester system is indicated in light of scheduling practices of related courses, especially Mathematics 11. How often do students take Chemistry 11 before they take Mathematics 11? Do girls and boys differ in the effect, if any, of scheduling Mathematics 11 and Chemistry 11 concurrently?

More precise methods of analysis and control of variables could determine whether there is a significant interaction between the sex of students and the variety of school programmes now being offered to students in high schools. Course enrolments show that little attempt is made to schedule courses consecutively or on the basis of related subject matter in the semestered high schools. The system was set up at least in part to make it easier for a student to acquire credits towards graduation (Nanaimo, 1968-69; Abbotsford, 1974); so perhaps it is the credit system rather than its offspring, the semester system, which requires re-examination. Coupled with the administrative problems of using one's staff effectively, it would appear that the acquiring of students' credits has become more important than the desire to teach more effectively. To do this would necessitate moving staff to other schools between semesters, as Saunders (1974) has pointed out. Failing this flexibility, courses are scheduled inevitably to fit the staff available.

3. The need for further examination of sex differences is the major outcome of this study. It is not enough merely to examine results on some achievement tests and then conclude that one sex achieves better than another. The many variables involved need to be identified and studied. Some other variables should be pursued in the area of teaching and learning styles and methods along the lines of the work of Welch (1969), Piaget (1959), and Lodge (1969). If it were to be found that girls and

boys do indeed differ in their reaction to school organization, to certain teaching methods and to male and female teachers, what would be the implications of such findings for future educational decisions, and by whom will such decisions be made?

REFERENCES

Abbotsford, Board of School Trustees. Report of the semester study committee, Abbotsford Senior Secondary School, December 1974.

Anderson, J.R. A multivariate analysis of attendance and achievement in secondary schools organized on traditional lines and semester systems. M.A. Thesis, University of Victoria, April 1972.

Ashmore, P.G. On teaching high school chemistry. *Science*, 1965, *148*, 1312-1314.

Bradley, J. & Hutchins, D. Concepts of science and scientists as factors affecting subject choice in secondary schools. *School Science Review*, 1973, *54* (190), 8-15.

Bridgham, R.G. Grading and enrolments in the sciences. *Science Teacher*, 1973, *40* (6), 41-42.

British Columbia, Department of Education. *Annual Report*. Victoria: Queen's Printer, 1968-69, 1969-70, 1970-71.

British Columbia, Department of Education. *Curriculum Guides* for Chemistry 11 and 12. Victoria, 1966.

British Columbia, Department of Education. Report to the Minister Donald L. Brothers of the Committee on School Utilization. (By J.L. Canty) Victoria, B.C., 1969.

British Columbia, Department of Education. Report on testing Chemistry 11 in June 1972 and January 1973. Research and Standards Division. (By C.B. Conway) Victoria, B.C., 1973.

British Columbia, Royal Commission on Education 1960. Chairman, N. Chant. *Report*. Victoria: Queen's Printer, 1960.

Coley, N.R. Prediction of success in general chemistry in a community college. *Journal of Chemical Education*, 1973, *50B*, 613-615.

Conway, C.B. Letter, 1976.

Coulson, E.H. Nuffield chemistry at "A" level. *Education in Chemistry*, 1973, *6*, 200-202.

Crawford, D.H. Relating mathematics and science at the secondary level. *Ontario Mathematics Gazette*, September 1971, pp. 1-16.

Crawley, D.M. The effect of sex differences and absence of instructor on student attitudes. Unpublished paper, University of Victoria, March 1975.

Denny, R. The mathematical skill test for chemistry. *Journal of Chemical Education*, 1971, *48*, 84-86.

Downing, J. & Thomson, D. Sex role stereotypes in learning to read. Unpublished paper, University of Victoria, 1975.

Downing, J. et al. A cross-national survey of cultural expectations and sex-role standards in reading. *Journal of Research in Reading* (in Press) for 1979.

Farrell, W.T. Growing up male in America. *1974 Yearbook, The World Book Encyclopaedia*. Chicago: Field Enterprises Educational Corporation, pp. 92-105.

Field, T.W. & Cropley, A.J. Cognitive style and science achievement. *Journal of Research in Science Teaching*, 1969, *6*, (1), 2-10.

Frankel, E. *A comparative study of achieving and under-achieving high school boys of high intellectual ability*. Research of science, U.S. Department of Health, Education and Welfare, July 1957-1959. (Abstract)

Friedan, B. *The Feminine Mystique*. New York: Norton, 1963.

Gallagher, J.J. & Korth, W.W. Test every senior project: Attitudes of seniors concerning science. Paper presented at the National Association for Research in Science Teaching Conference, February 1969, Pasadena, California.

Glass, G.V. & Stanley, J.C. *Statistical methods in education and psychology*. New Jersey: Prentice-Hall, 1970.

Gleadow, Norman. The semester system in Vancouver schools and its effects on curriculum. Evaluation Research services of Vancouver Board of School Trustees, January 1975.

Hill, P.L. Semestering and its impact on Ontario. *OSSTF Bulletin*, 1974, *54*, 199-203.

Ingle, R.B. & Shayer, M. Conceptual demands in Nuffield "O" level chemistry. *Education in Chemistry*, 1971, *8*, 182-183.

Inhelder, B. & Piaget, J. *The growth of logical thinking from childhood to adolescence: An essay on the construction of formal operational structures.* Trans. A. Parsons & S. Milgram. New York: Basic Books, 1958.

Johnstone, A.H. Evaluation of chemistry syllabuses in Scotland. *Studies in Science Education*, 1974, *1*, 21-49.

Johnstone, A.H., Morrison, T.I. & Sharpe, D.W.A. Topic difficulties in chemistry. *Education in Chemistry*, 1971, *8*, 212-213, 218.

Keeves, John. Differences between the sexes in mathematics and science courses. *International Review of Education XIX/1973/1* (Special Number - the Education of Women), UNESCO Institute for Education, 1973, pp. 47-63.

Lockwood, J.B. Mathematical processes needed in learning high school chemistry and high school physics. M.A. Thesis, Western Michigan College, 1957.

Lodge, F. Student rating of science skills. *Science Education*, 1969, *53* (5), 421-422.

Malm, L.E. (ed.) *Laboratory manual for chemistry: An experimental science.* San Francisco: Freeman, 1960, 1963.

MacLennan, L.K. Semestering and non-semestering: A trial marriage. *OSSTF Bulletin*, 1974, *54*, 105-109.

Martineau, H. *Society in America.* London: Saunders-Otley, 1837.

McLellan, A. (ed.) *Teachers guide for chemistry: An experimental science.* San Francisco: Freeman, 1963-70.

Merrill, R.J. & Ridgway, D.W. *The CHEM Study story.* San Francisco: Freeman, 1969.

Miller, L.C. Another view of CHEM Study. *The Science Teacher*, 1967, *34*, 60.

Nanaimo, Board of School Trustees. Semester plan: Nanaimo District Secondary School. Nanaimo, B.C., September 1968.

Nanaimo, Board of School Trustees. Preliminary evaluation of semester system at Nanaimo Disrict Secondary School. Nanaimo, B.C, July 1969.

Neter, J. & Wasserman, W. *Fundamental statistics for business and economics.* Boston: Allyn & Bacon, 1966.

Newbury, N.F. *The teaching of chemistry.* London: Heinemann, 1965, pp. 133-134.

Newcomer, M. *A century of higher education for women.* New York: Harper Bros., 1959.

New Statesman. New minds for the new world. September 8, 1956, p. 279.

Ostle, B. *Statistics in research.* Ames: Iowa State University Press, 1969.

Overall, J.E. & Klett, C.J. *Applied multivariate analysis.* New York: McGraw-Hill, 1972.

Pearson, E.S. & Hartley, H.O. Charts and the power function for analysis of variance tests derived from the non central F distribution. *Biometrika*, 1951, *38*, 112-130.

Pimental, G. (ed.) *Chemistry: An experimental science.* Prepared by Chemical Education Material Study (CHEM Study). San Francisco: Freeman, 1960-1963.

Preston, R.C. Reading achievement of German and American children. *School and society*, 1962, *90*, 350-354.

Rising, R.G. Developing ancillary skills: Mathematics deficiencies in the science classroom. *The Science Teacher*, 1967, *84*, 29-31.

Rothman, A.I., Welch, W.W. & Walberg, H.J. Physics teacher characteristics and student learning. *Journal of Research in Science Teaching*, 1969, *6*(1), 59-63.

Saunders, R.G. Semestering. *OSSTF Bulletin*, 1973, *53*, 203-206.

Shayer, M. CHEM Study in an English school: A personal view. *Education in Chemistry*, 1970, *7*, 54-55. (a)

Shayer, M. How to assess science courses. *Education in Chemistry*, 1970, *7*, 182-186. (b)

Shayer, M. and Wylam, H. The distribution of Piagetian stages of thinking in British Middle and Secondary school children. (II 14-16 year olds and sex differences). *British Journal of Educational Psychology*, 1976, *46*, 164-173.

Smeltz, J.R. Retention of learning in high school chemistry. *The Science Teacher*, 1956, *23*, 285-305.

Speed, F.M. & Hocking, R.R. The use of R notation with unbalanced data. *The American Statistician*, 1976, *30* (1), 30-33.

Stolee, Lief. Math A's increase 1391.7% in post exam schools. *St. John's Edmonton Report*, June 2nd, 1978.

Vancouver, Board of School Trustees. An evaluation of semestering in Vancouver secondary schools during 1970-1971 school year. Department of Planning and Evaluation Research Report No. 71-19. (By Allan G. Moodie) Vancouver, 1971.

Victoria, Board of School Trustees. Semester system: A preliminary evaluation. The Research and Testing Division, School District 61, Greater Victoria, 1970.

Welch, W.W. Correlates of course satisfaction in high school physics. *Journal of Research in Science Teaching*, 1969, *6*, 54-58.

Williams, H.R. & Koflsche, C.L. Organization of chemistry classes: A comparative study of two patterns of teaching of high school chemistry. *The Science Teacher*, 1967, *34*, 52-54.

Winer, B.J. *Statistical principles of experimental design.* New York: McGraw-Hill, 1962.

APPENDIX A

Test Methods of Distribution
Sampling Procedures

ERRATA

Test Forms

A number of errors were present in all three forms as follows:

Form A. Question 1. Both the fourth and fifth distractors were numbered (4) and the word "head" was omitted in distractor (2). Both these errors were corrected in the papers sent out in January 1973. This question was incorrectly keyed for (2) as the correct answer instead of (3) throughout the study in the departmental grading. This was corrected in the regrading of the samples by computer.

Form B. In the preamble to questions 10 and 11, the word "obtain" was omitted making the question much less intelligible. In question 11 the word "oxide" had been incorrectly inserted. In question 27 an arrow had been omitted from an equation. None of these errors were corrected in the January 1973 distribution.

Form C. In Item 27, FeS_2 had been incorrectly written as Fe_2S. This was corrected in the January forms.

Research and Standards Branch
Department of Education, Victoria

Directions for Administering

B.C. CHEMISTRY 11 TESTS
Forms A, B, C

June 1972

Purpose of the Survey

These tests are general achievement tests at the Chemistry 11 level and are designed to measure the extent of a student's mastery of the content and applications of Chemistry 11 material. Losses and gains in semestered and non-semestered schools will be compared. Each student will receive one of the three forms which are to be treated as one for administrative purposes. After the norms have been calculated the tests will be available to teachers of Chemistry students.

Examiner's Preparation

The examiner should become thoroughly familiar with the directions for administration that follow. BE SURE THAT THE NAME AT THE TOP OF THIS PAGE CORRESPONDS TO THE NAME OF THE TESTS YOU ARE ABOUT TO ADMINISTER. It is best to read the directions through twice before attempting to administer the examination. Many of the questions that examinees will ask can be anticipated and problems of organization can be foreseen and prepared for. Speed in getting the pupils down to work is important. Candidates should be allowed only a minimum of time to fill in the blanks on the cover of the booklet. As soon as a majority has finished, the slower ones should be told to complete the information after the end of the examination. The test papers should be checked carefully to see that they have done so.

Chemistry 11 classes may be combined for testing, providing that students are not crowded together. Each student should have at least three feet of desk space so that he can work comfortably. When a large group is to be tested, proctors may be needed to help distribute the tests. There should be one proctor for every 10 or 20 examinees. The proctors should understand in advance what they are to do so that the booklets may be distributed quickly and collected without confusion.

Students who are to be tested

All students who are taking or who have taken Chemistry 11 this school year are to be tested, i.e. include students who completed Chemistry 11 in the fall semester (end of January) and are still in school. The system under which each pupil took Chemistry 11 will be recorded on the cover of each test.

Each pupil should have a pen and two pencils and a sheet of scratch paper.

Class Record Sheets

Class record sheets are not required. They will be prepared in Victoria. But a list of absentees is necessary and space is provided for this on the final page of these directions. In order to account for all papers written, a count of the number of students writing also must be made.

Timing

The actual working time of the test is 60 minutes in one continuous session, but an additional 5 minutes may be required for the distribution of tests and the completion of the cover page. The examiner should have a watch with an easily read second hand and should pay particular attention to the time so that he does not forget to issue the "Stop" instruction exactly when it is required. The Time Record Sheet which follows these directions should be completed as soon as the students have begun and should be corrected afterwards, if necessary, to indicate any error in timing.

Distribution of test papers

The examiner probably will have noticed that three forms (A, B, C) of the Chemistry 11 tests are alternated in the package. This will ensure random sampling throughout the school and throughout the province. Tests are to be distributed as they are received; the distribution should be made as if there were only one form.

Directions to the Pupils

(1) When the class is seated and the proctors or the pupils in the front row are assisting in passing out the booklets, the examiner should say:

"THIS IS A TEST TO MEASURE YOUR ACHIEVEMENT IN CHEMISTRY 11. THE SCORES ARE TO BE PLACED ON YOUR PERMANENT RECORD AND WILL BE COMPARED WITH THOSE OF THE OTHER CHEMISTRY 11 STUDENTS IN THE PROVINCE.

"FILL IN THE BLANKS ON THE COVER OF THE BOOKLET BUT DO NOT OPEN THE TEST UNTIL YOU ARE TOLD TO DO SO. PRINT YOUR SURNAME FIRST, AND THEN PRINT YOUR INITIALS IN BLOCK CAPITALS. THE NUMBER OF THIS SCHOOL DISTRICT IS .... THE DATE IS .... BE SURE TO INDICATE THE PROGRAMME YOU ARE TAKING THIS YEAR, THE SYSTEM UNDER WHICH YOUR CHEMISTRY 11 COURSE IS BEING TAUGHT, AND, IF ON THE SEMESTER SYSTEM, THE SEMESTER YOU TOOK OR ARE TAKING CHEMISTRY 11.

"WHEN YOU HAVE FILLED IN THE BLANKS, READ THE GENERAL DIRECTIONS BUT DO NOT TURN THE PAGE UNTIL YOU ARE TOLD TO DO SO."

(2) Allow a reasonable time for filling in the spaces on the cover page and reading the General Directions. When most of the students have done this, the examiner should say:

"IF YOU NEED EXTRA SPACE FOR CALCULATIONS USE THE SCRATCH PAPER. BUT ONLY THE ANSWERS YOU PUT ON THE TEST BOOKLET WILL BE MARKED.

"ARE THERE ANY QUESTIONS? NO QUESTIONS MAY BE ASKED AFTER THE TEST BEGINS." (Students who have not completed the cover page may be told to do so at the end of the test.)

(3) When all are ready, say:

"YOU MAY FIND IT DIFFICULT TO COMPLETE THE TEST IN THE TIME THAT HAS BEEN ALLOTTED. BUT DO NOT BE HASTY. READ EACH QUESTION CAREFULLY. DO THE BEST YOU CAN. WORK RIGHT THROUGH TO THE END AND THEN GO BACK TO TRY QUESTIONS YOU HAVE MISSED. OPEN THE BOOKLET AND BEGIN THE TEST."

(4) Note the exact time when you say "Begin" and write it on the Time Record that follows. Add 60 minutes and make the entries on the Time Record. Do not allow the students any extra time for reading directions that may appear in the body of the test.

(5) At the end of 60 minutes, say:

"STOP. EVEN IF YOU HAVE NOT FINISHED, YOU MUST CLOSE YOUR BOOKLETS. CHECK THE INFORMATION THAT YOU HAVE FILLED IN ON THE COVER."

(6) Have the booklets collected at once. Make sure that pupils have entered all information necessary for identification and classification. It is especially important that the students' surnames are printed legibly, followed by their initials.

(7) Complete the List of Absentees (next page). Arrange the booklets in alphabetic order and place this Time Record and the List of Absentees on top. See that the total number of booklets corresponds to the number of pupils who have written the test. Under no circumstances should any attempt be made to mark the tests. Unused booklets should be separated from the used booklets by means of a sheet of paper or separate tying, and be enclosed in the same parcel. Do <u>not</u> attempt to sort out Forms A, B, C.

Please return immediately:

(1) Achievement Tests (Forms A, B, C mixed), Time Record and List of Absentees.

(2) All unused tests (as in Paragraph 7 above).

to: Research and Standards Branch,
Department of Education,
Parliament Buildings,
Victoria, B.C.

TIME RECORD

B.C. Chemistry 11 Tests - Forms A, B, C

(to be returned with the test papers)

	Hr.	Min.	Sec.
End of preliminary directions: ("Open the booklet and begin the test.")			
Add	--	60	--
Time to say: "Stop. .... Close your booklets."			

(If, for any reason beyond your control, timing is not strictly adhered to, enter the actual times and write an explanatory note on this sheet so that a decision may be made as to whether the scores shou be included in the provincial norms.)

This number is the total number of students who wrote one form of the Chemistry 11 tests (i.e. equal to the number of used booklets).

NUMBER OF STUDENTS WRITING ....

LIST OF ABSENTEES

List the names of all students still in school who did not write any form of the Chemistry 11 test. (Surname, followed by initials, and M or F for sex.)

	Sex		Sex
....	...	....	...
....	...	....	...
....	...	....	...
....	...	....	...
....	...	....	...
....	...	....	...

....
Administrator of Test

BRITISH COLUMBIA TESTS

CHEMISTRY II

FORM A

Research and Standards Branch,
Department of Education, Victoria, British Columbia.

Fill in the blanks below but *do not open the booklet* until you are told to do so.

ame________________________ Age________ (Last birthday) Sex________ (M or F)

ate________________________ Grade________ Home-room Div. No.________

:hool________________________ School District No.________

rogramme: Academic Technical________ Commercial________ Comm. Services________ Industrial________

:heck (✓) propriate e) Visual and Performing Arts________ Agriculture________ Other programmes for particular occupations________

Have you taken or are you enrolled in the following courses?

	Presently Enrolled	Previously Taken		Presently Enrolled	Previously Taken
iology 11			Mathematics 10		
hemistry 11 1st semester			Mathematics 11		
2nd semester			Mathematics 12		
10-month course			Physics 11		
			Science 10		

;eneral Directions: The purpose of this test is to measure your knowledge of the prescribed course in Chemistry 11. You will have 60 minutes for the test. Do not spend too much time on any one item; answer easier questions first, then return to the difficult ones if you have time. A Periodic Table of the elements is printed on the back page. It may be used whenever an atomic weight is needed. Use approximate rounded values (e.g., manganese (Mn), atomic weight 55). Calculations and rough work may be done on extra paper but only the answers placed in the correct answer spaces will be marked. They may be either digits or capital letters. Wild guessing of multiple-choice answers is sure to result in a low score, so study the items carefully. Now wait until you are told to begin and then turn the page and proceed.

Printed by K. M. MacDonald, Printer to the Queen's Most Excellent Majesty
in right of the Province of British Columbia.
1972

BRITISH COLUMBIA TESTS

CHEMISTRY 11

FORM B

Research and Standards Branch,
Department of Education, Victoria, British Columbia.

Fill in the blanks below but *do not open the booklet* until you are told to do so.

Name____________________________ Age________ (Last birthday) Sex________ (M or F)

Date________________________ Grade________ Home-room Div. No.________

School____________________________ School District No.________

Programme: Academic Technical________ Commercial________ Comm. Services________ Industrial________

(Check (✓) appropriate one) Visual and Performing Arts________ Agriculture________ Other programmes for particular occupations________

Have you taken or are you enrolled in the following courses?

	Presently Enrolled	Previously Taken		Presently Enrolled	Previously Taken
Biology 11			Mathematics 10		
Chemistry 11 1st semester			Mathematics 11		
2nd semester			Mathematics 12		
10-month course			Physics 11		
			Science 10		

General Directions: The purpose of this test is to measure your knowledge of the prescribed course in Chemistry 11. You will have 60 minutes for the test. Do not spend too much time on any one item; answer easier questions first, then return to the difficult ones if you have time. A Periodic Table of the elements is printed on the back page. It may be used whenever an atomic weight is needed. Use approximate rounded values (e.g., manganese (Mn), atomic weight 55). Calculations and rough work may be done on extra paper but only the answers placed in the correct answer spaces will be marked. They may be either digits or capital letters. Wild guessing of multiple-choice answers is sure to result in a low score, so study the items carefully. Now wait until you are told to begin and then turn the page and proceed.

Printed by K. M. MacDonald, Printer to the Queen's Most Excellent Majesty
in right of the Province of British Columbia.
1972

BRITISH COLUMBIA TESTS

CHEMISTRY II

FORM C

Research and Standards Branch,
Department of Education, Victoria, British Columbia.

Fill in the blanks below but *do not open the booklet* until you are told to do so.

Name.. Age.................. (Last birthday) Sex.................. (M or F)

Date.. Grade.............. Home-room Div. No. ....

School.. School District No. ....

Programme: Academic Technical.................. Commercial.................. Comm. Services.................. Industrial..................

(Check (√) appropriate one) Visual and Performing Arts.................. Agriculture .... Other programmes for particular occupations..................

Have you taken or are you enrolled in the following courses?

	Presently Enrolled	Previously Taken		Presently Enrolled	Previously Taken
Biology 11			Mathematics 10		
Chemistry 11 1st semester			Mathematics 11		
2nd semester			Mathematics 12		
10-month course			Physics 11		
			Science 10		

General Directions: The purpose of this test is to measure your knowledge of the prescribed course in Chemistry 11. You will have 60 minutes for the test. Do not spend too much time on any one item; answer easier questions first, then return to the difficult ones if you have time. A Periodic Table of the elements is printed on the back page. It may be used whenever an atomic weight is needed. Use approximate rounded values (e.g., manganese (Mn), atomic weight 55). Calculations and rough work may be done on extra paper but only the answers placed in the correct answer spaces will be marked. They may be either digits or capital letters. Wild guessing of multiple-choice answers is sure to result in a low score, so study the items carefully. Now wait until you are told to begin and then turn the page and proceed.

Printed by K. M. MacDonald, Printer to the Queen's Most Excellent Majesty
in right of the Province of British Columbia.
1972

LETTER FROM DR. CONWAY

February 1976

Mrs. Lavinia Greenwood,
687 Mt. Joy Avenue,
Victoria, B.C.

Dear Mrs. Greenwood:

In reply to your request for an explanation of our method of sampling, here are the steps that were taken:

1. For administration of the tests the "sample" was 100% of those attempting to complete the course in the year or semester being tested, i.e. it was not a sample; it was a total population. Such a population will never be equal to the reported enrolment, however. 'Enrolment' is the maximum number of names listed for the course during the time period under consideration. It will include:
   (a) those ill or absent for any other reason on the date of testing,
   (b) those who have moved (a tremendous number in total),
   (c) those who have dropped out.

2. The chief factor in obtaining reliable measurement is the length of the measuring instrument, i.e. the number of markable items. Here we are faced with a dilemma: the time of administration must be short enough to fit the shortest time period devoted to the subject in all of the different timetables which may allot 40, 50, or even 120 minutes in different schools. If too many items are prescribed for the short time period, it becomes a test of speed rather than power.

This problem is solved by breaking the test into a number of "equivalent" forms, each one short enough to be completed in the prescribed time period. ("Equivalence" is of course only approximate and we have often found that forms of standardized tests that have sold hundreds of thousands of copies were not what we would call "equivalent" when administered to 20,000 or more pupils.)

In order to determine the "equivalence" of items we have to administer them to "equivalent" populations or to administer all forms in succession to all pupils. (Even if we did so, we would have to randomize the administration of A, B and C because motivation on the third day might not be as great as on the first.)

Instead, we shipped the tests interleaved with Forms A, B, C, A, B, etc. in order. Instructions were to distribute these down the rows of pupils in the order in which they were received, in order to obtain three equivalent samples of the total population. Although B or C forms would be on top of the pile with equal frequency, we found a slight tendency for A's to be administered with greater frequency; the teachers buried form B or C and there were more of the latter left over when the number being tested was not divisible by three. The meticulous mind dictates that Form A must always be given to the first pupil in the left-hand row!

I should mention that we have found this method of randomization to be vastly superior to any method of attempting to obtain samples of "equivalent schools" or "equivalent classes". We attempted for several years to obtain samples based on a selection of schools in order to proceed with the scaleing of examinations before marking was completed. We never succeeded. Even carefully selected halves of all the schools in the province, selected by hindsight on the basis of scores in Mathematics, would prove completely unequal in English.

3. Although it is quite possible that in a small class, say of 9 pupils, the three best pupils wrote Form A and the three worst ones wrote Form C, I think that province-wide the sample populations would be equivalent.

Our procedure was then as follows:

(a) File the papers as received according to school district, school and class.

(b) Split this arrangement into 3 files, one for each of the forms which are now separated for marking.

(c) Mark the tests, alphabetize them and type separate lists for each class for each form. This is necessary because the forms may not be "equivalent" and the scores may have to be compared with different norms.

4. One objection to machine scoring and computer handling of scores is the inability of any mechanical device to distinguish between a zero or very low score and a "D.N.W." The latter is defined for marking purposes as any pupil who, although present and completing the cover of the test, obviously did not make a conscientious effort to complete as many items as possible to the best of his ability. Examples are pupils who correctly or incorrectly completed only one or two items in half an hour, or ones who answer multiple-choice items 1, 2, 3, 4, 5, 1, 2, 3, 4, 5, etc. throughout the test.

Such scores are supposed to be excluded from any sampling or statistical analysis because although they are few in number, they can greatly distort mathematical data. Occasionally our hand-marking and hand-sampling methods permit one or two of these to slip in, but they are supposed to be removed.

5. It has been our procedure to conduct item-analysis and reliability studies on samples of 300, whether the number of pupils tested was 1500 or 45,000. There are three reasons: (1) The largest sheets of squared

paper we could conveniently use for tabulation had about 110 squares and allowing for headings, totals, half-scores, etc. we filled three sheets with the upper, middle and lower parts of the distribution. (2) In studies we made years ago on the reliability of item-validities we found that samples of 100 or so were quite likely to be unreliable but that by the time 300 was reached we were almost as high on the curve as if we had used 600 or 1000. (3) Per cents of each third can be obtained without calculation.

Moreover, in test construction, statistical data can only be a rought guide. Why calculate to three or four significant decimals if you are only using the first one as an approximation? Curricular validity is the prime consideration, and we have frequently had to discard items of high statistical validity in order to obtain balanced coverage of the course.

The samples of 300 were obtained as follows from the tests which were in alpha order by class, by school and by school district with the DNW's presumably removed. If there were 1567 of these papers we needed approximately one-fifth. Five papers squeezed together were approximately one-quarter inch. Therefore going down through the piles an estimated quarter-inch (no counting) and turning them over, individual papers were removed. These now made up a pile about 15 inches deep and when counted carefully, it was found to contain 319 papers. Nineteen were removed at random (about one in seven-eighths of an inch) and returned to the population. If there had been a deficiency, we would have gone back to the total and say, removed one additional paper every 2-1/2 feet.

The selection is completely 'blind', i.e. there is no reference to school or score. Only when the sample is finalized at 300 are they sorted according to score, after which they are analyzed. We would expect to have about twice as many papers from Vancouver as from Surrey or Greater Victoria, but we make no attempt to modify the sample to ensure that there are. Nevertheless, with the exception of the DNW's we are sure that the 300 is representative of the province of B.C.

Yours very truly,

C.B. CONWAY
Ex-Director,
Research and Standards.

APPENDIX B

Extra Tables, Definitions, Mathematical Assumptions and Supporting Statistical Theory And Data For the Analysis of Variance

TABLE VI

Variances of the Mean Scores and

Homogeneity of Variance Data (χ^2)

Form		Prog R	Prog S	Prog F	Total	χ^2	p*
	All	28.57	28.04	27.57	28.20	0.09	0.95
A	Boys	32.60	29.59	28.21			
	Girls	19.15	23.90	26.06			
	All	19.57	17.70	18.60	18.56	0.75	0.69
B	Boys	20.33	17.57	20.33			
	Girls	17.56	17.56	15.36			
	All	24.93	29.23	22.14	23.35	5.84**	0.054
C	Boys	27.81	32.53	23.64			
	Girls	20.77	24.35	19.84			

*p = probability that the variances are equal.

**significant χ^2

TABLE VII

Kuder Richardson Reliability Coefficients

Boys and Girls

	Form A			Form B			Form C		
Prog.	Boys	Girls	Pooled	Boys	Girls	Pooled	Boys	Girls	Pooled
R	.816	.716	.793	.698	.668	.689	.775	.705	.752
S	.797	.758	.785	.666	.690	.675	.739	.763	.790
F	.797	.783	.790	.718	.635	.686	.739	.711	.726

TABLE VIII

Incompleted Tests

Form	Item #	Semester R			Semester S			Semester F			Totals
		Boys	Girls	Total	Boys	Girls	Total	Boys	Girls	Total	
A	23					1	1				1
	24					1	1				1
	25					1	1	1		1	2
	26	1	1	2	2	4	6	4	4	5	13
	27	1	2	3	3	4	7	6	3	9	19
	28	3	4	7	3	4	7	6	8	14	28
	29	17	15	32	13	21	34	31	19	50	116
	30	22	22	44	22	37	59	49	28	77	180
B	26				1		1				1
	27				1		1				1
	28	3	1	4	3	2	5	1	2	3	12
	29	3	1	4	3	2	5	1	4	5	14
	30	4	1	5	5	4	9	3	4	7	21
C	23	1		1							1
	24	2		2							2
	25	2		2	1		1				3
	26	2		2	1	1	2		1	1	5
	27	2	1	3	2	1	3		3	3	9
	28	3	2	5	3	8	11	5	8	13	29
	29	3	3	6	3	9	12	7	10	17	35
	30	9	16	25	10	23	33	16	27	43	101

TABLE IX

Mean Scores by Mathematics Groups: Semester 4, January 1973

Math Group		Form A	Form B	Form C	Combined
Math 1	Population	12.65*	13.30	13.37	13.10
	Sample	13.32	13.78	13.43	13.51
Math 2	Population	13.45*	13.40	13.26	13.37
	Sample	13.34	13.78	13.47	13.54
Math 3	Population	11.72*	12.75	13.09	12.50
	Sample	12.87	12.35	13.43	12.94

The above table was subjected to the χ^2 test for goodness of fit of the sample means to the population means distributions using the following formula (from Bernard Ostle, *Statistics in Research* [Iowa State University Press, 1969], p. 31):

$$\chi^2 = \sum (O - E)^2/E$$

where E = population means and O = sample means

$$_{p}\chi^2_{(4)} = 0.20225 < {}_{.005}\chi^2_{(4)} = 0.207$$

(see Ostle, tables, p. 523).

Therefore "sample" distribution matches the population distribution to within one-half of one percent. Therefore there is no "sampling bias" when the scores are cross-classified by mathematical groupings.

*Item 1 wrongly keyed.

TABLE X

Mean Scores for Populations and Samples

Semester		Form A	Form B	Form C	Combined
(10 month)	Population	13.33*	13.54	13.79	13.52
1	Sample	14.32	13.40	13.66	13.74
2	Population	13.28*	13.57	13.58	13.48
	Sample	14.13	13.59	13.63	13.74
3	Population	11.60*	11.32	12.07	11.83
	Sample	12.83	12.15	12.47	12.48
4	Population	12.58*	13.18	13.26	13.01
	Sample	13.21	13.45	13.44	13.37
Total	Population	12.70	12.97	13.18	12.95
	Sample	13.62	13.15	13.30	13.34

* Item 1 wrongly keyed.

Source: Population means, C. B. Conway, 1974.
Sample, ANOVA 15B after rescoring.

Maximum score = 30

Definitions

The design of this study is sometimes referred to as quasi-experimental being more in the nature of a survey than an experiment.

The following definitions are included in order to clarify the methods used to control the major variables which were expected to influence the test scores.

Surveys. "A survey is the collection of data about characteristics of existing things, with no special control over any of the factors influencing the variable of interest" (Neter & Wasserman, 1966, p. 10).

Experiments. "An experiment is used by statisticians to refer to the collection of data when control is exercised over one or more factors, to determine the influence of these on the variable of interest" (Neter & Wasserman, 1966, pp. 10-11).

Cross-classification. "Classification entails the setting up of groupings or classes by classification methods for the alternative outcomes of a variable" (Neter & Wasserman, 1966, p. 39). In surveys, since there is no way to exercise controls over factors which might bias the data, the technique of cross-classification can control variables and bring out in the analysis their possible effects upon the dependent variable (Neter & Wasserman, 1966).

The classification categories used were: (1) sex; (2) programmes; and, (3) mathematics group.

Programmes. This refers to the school programme - whether it was a regular ten-month year (Programme R) or a semester system. If the school took Chemistry 11 in a semestered school the first and second semesters were separated into Programmes F and S respectively.

Mathematics Group. Students were later subdivided into three groups according to whether they had taken Math 11 with Chemistry 11 (Group 1) before (Group 2) or after (Group 3).

A Mathematical Model for the Analysis of Variance And the Assumptions Necessary for Interpretation of the Results

In undertaking an analysis of variance on data with unequal cell frequencies, more than one method is open to the researcher. Different methods test different hypotheses, but all the methods are based on the General Linear model, and use a least-squares solution (Overall & Klett, 1972). The model may be expressed as follows in a two-way classification:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk}$$

Where μ is the overall mean

α_i is a constant for the i^{th} sex

β_j is a constant for the j^{th} programme

$\alpha\beta_{ij}$ is the effect of interaction of the sex with the programmes

ε_{ijk} is the residual error associated with the k^{th} individual whose score is Y_{ijk} in the i^{th} sex, j^{th} programme.

To obtain a solution, a number of restrictions must be imposed on the equation. The interpretation of the solution will depend upon the set of restrictions which are chosen.

Assumptions

For an analysis of variance to be valid, a number of assumptions must be met regarding the form of the data and the methods of sampling (Glass & Stanley, 1970, p. 358, 421, 424; Overall & Klett, 1972, p. 431). These assumptions are those of a normal distribution of the functions and are as follows:

(1) The scores were sampled at random from

(2) populations showing a normal distribution of the scores.

(3) The variances of the populations were equal, i.e., $\sigma_{1.}^2 = \sigma_{1.}^2 = \sigma_{2.}^2$ and $\sigma_{.1}^2 = \sigma_{.2}^2 = \sigma_{.3}^2$

(4) The samples were independent.

(5) The linear model is appropriate for these data.

The assumptions are discussed below as they relate to the data for this study.

Assumption 1 can be said to have been met by the method of systematic stratified sampling previously described.

Assumption 2. The population used was the same population as was studied by Dr. Conway. Figures 3 and 4 illustrate the frequency distribution of the population and samples respectively of one programme. Both population and sample may be considered large enough to satisfy this requirement (Tables I and II).

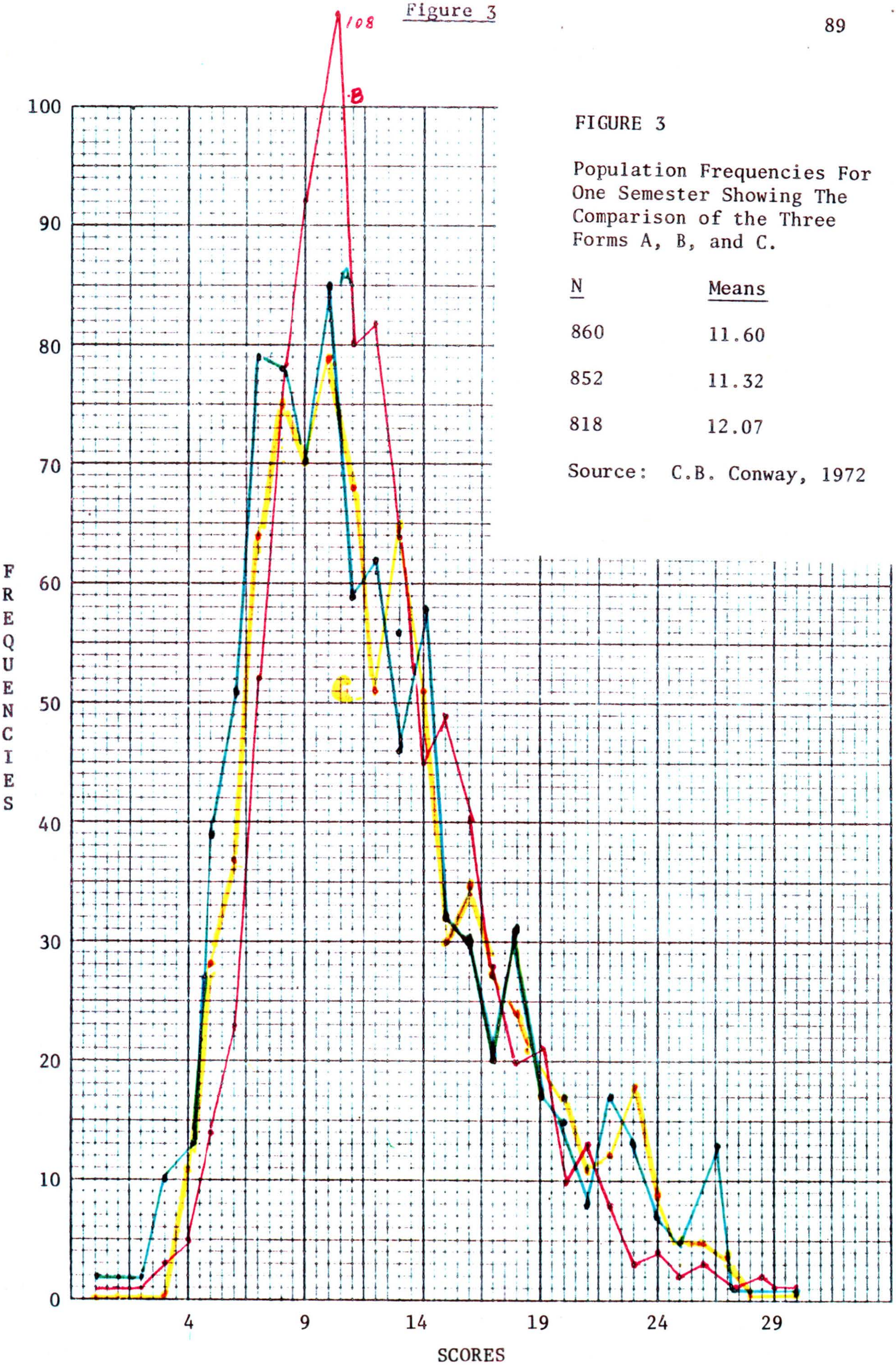


FIGURE 3

Population Frequencies For One Semester Showing The Comparison of the Three Forms A, B, and C.

N	Means
860	11.60
852	11.32
818	12.07

Source: C.B. Conway, 1972

Figure 4

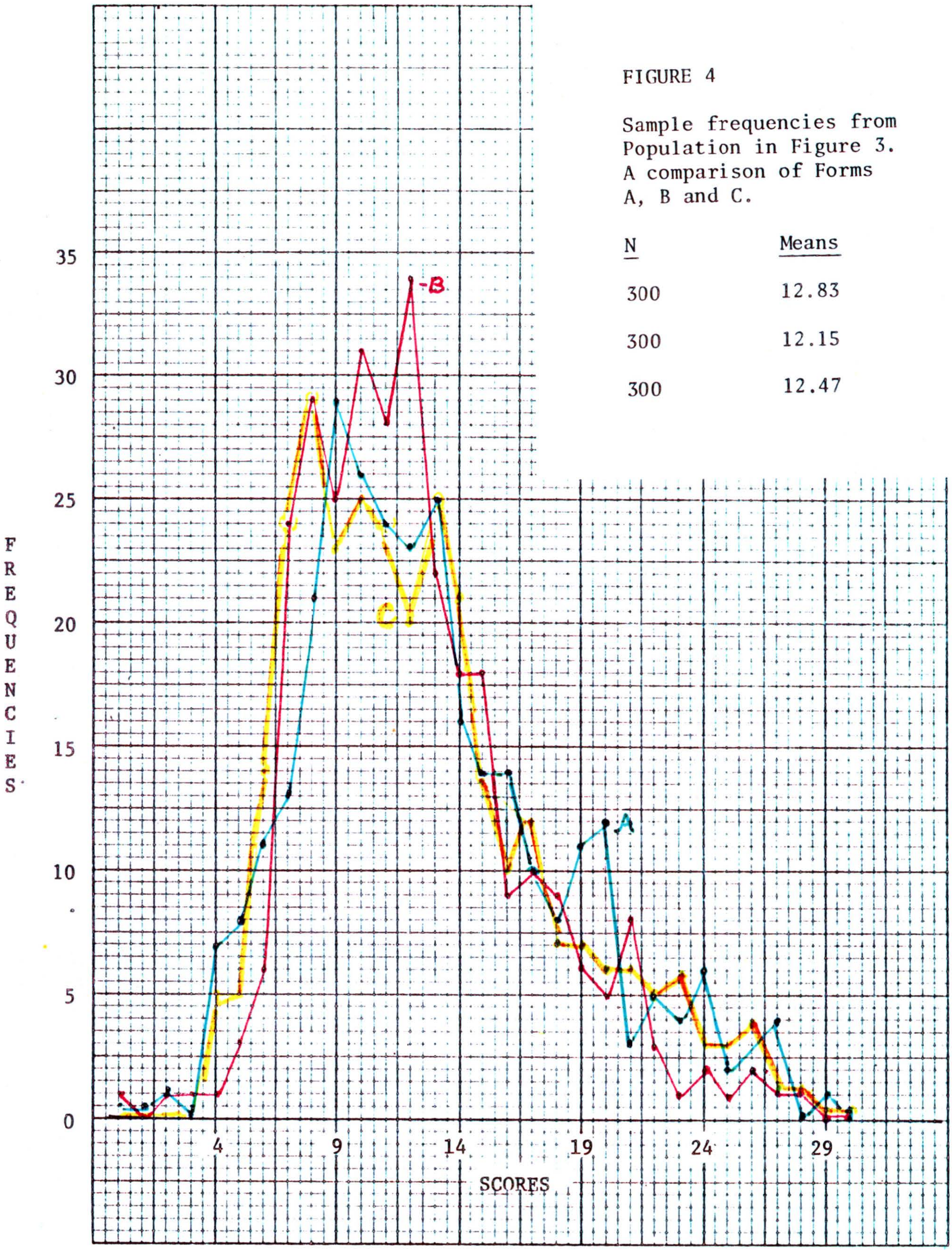


FIGURE 4

Sample frequencies from Population in Figure 3. A comparison of Forms A, B and C.

N	Means
300	12.83
300	12.15
300	12.47

Assumption 3. The presence of significant differences in variances of the three forms was one reason for treating these three forms separately and not pooling them as equivalent. Differences across the three forms were greater than differences between programmes for any one form. Girls showed consistently lower variances than boys (Tables IV and VI). Variances across programmes with the sexes pooled could be rated as homogeneous for Forms A and B, but was barely acceptable for Form C (Table VI).

Assumption 4 is strictly met by the nature of the design, i.e., by the use of exclusive categories.

Assumption 5. The assumption of a linear model is the only logical one, since only categorical variables are involved. However, the model is less than ideal and shows high residuals.

Level of Significance

The power of the *F* test was calculated in order to decide the appropriate level of significance to adopt. The following formula was used (Glass & Stanley, 1970):

$$\text{Power} = \emptyset = \sum_{j=1}^{J} \frac{nj(\mu j - \mu)^2}{J\sigma^2}$$

where $\mu = (\mu_1 - \mu_2 - \mu_3 - \mu_4)/J$

μj = mean for each semester

J = 4 (number of semesters)

σ^2 = population variance (estimated from the grand mean) = 22.6018

nj = 300 (sample size).

The results are presented below in Table X.

TABLE XI

Power Table

	Power	$\alpha = .05$		$\alpha = .01$	
	$\emptyset$	$1 - \beta$	β	$1 - \beta$	β
Form A	2.2666	0.97	.03	.93	.07
Form B	2.0885	0.95	.05	.86	.14
Form C	18374	0.89	.11	.72	.28

Since α = probability of making a Type I error, and β = probability of making a Type II error, it was concluded that the selection of $\alpha = .05$ gave the most reasonably balanced choice (Pearson & Hartley, 1951; Glass & Stanley, p. 357).

The Computer Programmes Used

The Analysis of Variance was carried out with the aid of the Computer Programme entitled ANOVA 22, from the University of Victoria's computer programme library in the Faculty of Education. The computations in the programme are based upon Winer (1962, pp. -24-227, 291-297) but it is not known which restrictions on page 224 are applied. It is now recognized that the final sums of squares in cases of unequal cell frequencies will be dependent upon the choice of restrictions (Speed & Hocking, 1976), and these need to be stated in the literature of prepared programmes. Therefore, the results from ANOVA 22 do not necessarily agree with results from other programmes which may be using different assumptions and different methods of computation.

TESTAN. This programme was used for grading the tests and computing test statistics with an item analysis. (Faculty of Education Programme Library, University of Victoria).

APPENDIX C

History, Evaluative Comments, and Conceptual Analysis of CHEM: Study

The History of the CHEM Study Programme

As the result of the work of a ten-man revision committee, the Department of Education selected a new chemistry course for the B.C. schools in 1966. They selected as the basic texts those of the CHEM Study programme, which was at that time in use in many western parts of the United States. Originally designed to be a one-year course, B.C. made it a two-year sequence, calling the two parts Chemistry 11 and Chemistry 12, and making some further slight modifications in emphasis.

The original writers of the course made their own follow-up study in 1969, and published their report in the form of a book called *The CHEM Study Story* (Merrill & Ridgway, 1969).* The following history of the course and comments on aims, objectives and achievements are taken extensively from this book.

CHEM Study stands for Chemical Education Material Study. It is a programme initiated in the United States in 1959 and supported by the National Science Foundation with grants totalling $2,800,000 up to 1968 to Berkeley and to Harvey Mudd College in California (p. 1).

---

*All bracketed page numbers refer to this book unless otherwise indicated.

Sparked perhaps by the shock of Sputnik that sent thousands of Americans to question the efficiency of their schools, a steering committee of the American Chemical Society became convinced that old chemistry courses commonly being offered in high schools were seriously deficient. So a planning committee was set up containing mainly university professors and some high school teachers from urban schools. No private or independent schools were represented nor university professors of education or curriculum specialists.

Six writing teams each undertook a different set of chapters. The sections were then studied, and put together and extensively revised by the whole committee. More revisions were later undertaken after trials in 25 selected high schools over the first three years. The Educational Testing Service prepared and distributed multiple choice, open book achievement tests. A "Teachers' Guide" was developed and films made and distributed. By 1963, 550 schools had tried it, and final revisions were made in the fall of that year.

The book has since been translated into 13 languages, including an unauthorized Russian version. The latter's preface states that the book is distinguished for its teachability, rigour and orderly presentation (p. 54).

The goals of the course. The steering committee's chief criticism of the existing courses in the 1950's were threefold: (a) too much descriptive detail needing memorization and outdated technology; (b) inadequate emphasis on modern unifying concepts; (c) "cook book" laboratory work was thought to be "misinforming students about the nature of scientific investigations." These remarks may be considered equally valid for British Columbia's old chemistry course (Chemistry 91).

The goals of the new, however, were broad and suggestive perhaps of running after too many hares. CHEM Study's own four goals are summarized as follows:

> 1. To diminish the separation between scientists and teachers in understanding of science.
>
> 2. To encourage teachers to undertake further study in chemistry - to keep pace . . . and thereby improve their teaching methods.
>
> 3. To stimulate students to a profession in chemistry.
>
> 4. To instil in non-continuing students an understanding of the importance of science in current and future human activities (pp. 1-2).

More specifically, the textbooks emphasized the discovery approach by presenting laboratory investigations on a tape before it was to be discussed in class. The hope was to instil a keen observation, and a questioning attitude in students who will not be content with blind acceptance of dogmatic assertions.

The theme of the "Teachers' Guide" is *support*. Its audience is defined as "a lone teacher isolated from other chemistry teachers with somewhat less than ideal training in poorly equipped school" (McClellan, 1963, p. v0. The guide then instructs the teacher to deal with the first six chapters in a general way without taking up any of the ideas in detail. They specifically state that "a complete understanding or mastery is *not* aimed at" (p. 1), suggestive of the acquaintanceship philosophy.

British Columbia's goals. The British Columbia Department of Education created its own goals. The Chemistry 11 Curriculum Guide (1966) reiterated the emphasis on a laboratory centred course; but in dividing the course into two parts, it stressed that Chemistry 12 demands good laboratory equipment and a well prepared teacher. The general approach to the course is to be an effort to treat chemistry as a language - a language with the elements as an alphabet, formulas for words, and equations for sentences. Its philosophy is to place less emphasis on training chemists and yet to recognize that chemistry is a frequent prerequisite for other studies and careers. It is hoped to introduce the average student to the language of science and the scientific approach to problems as part of their general education.

Assessments of the Course

In their follow-up study entitled *The CHEM Study Story*, Merrill and Ridgway (1969) reported on their statistical surveys carried out between the years 1961-1965.

The study was conducted through information from ten universities and 7,500 students' records. Of these, 96 per cent were CHEM Study students and only 4 per cent trained in other ways; so one can fairly ask whether the proportion of controls was an adequate comparison. As in the University of Victoria, there were usually two courses open to students, a selective course for those qualifying and a regular course for the rest. Many variables were studied. It was found that the interval between high school and college, the size of the high school classes, time spent in the high school laboratory, and whether or not the student took high school physics did not significantly influence scores. These findings may come as a surprise to many people.

Scores on SCAT (School and College Ability Tests) and CHEM Study achievement tests in 1962 showed no significant correlation, and CHEM Study students tended to do poorly also in CEEB examinations (College Entrance Examination Board). It was estimated that the potentially best students were most handicapped in this latter examination by a

CHEM Study preparation in competition with the traditionally trained. Follow-up studies compared students' GPA's in their freshman college chemistry courses. They found no difference in the performance of former CHEM Study students who had enrolled in the regular college course; but in the selected groups, it was claimed that CHEM Study had been a "distinct advantage". This was a disappointment since better preparation for college for *all* students had been one of the original stated goals of the project.

An attempt to assess the broad general aims in part four of their stated goals was abandoned as too difficult. Fifteen testing agencies were approached without success.

Independent Opinions

Early enthusiastic comments gradually gave way to the cautious criticism of experience. In 1965, P.G. Ashmore compared the CHEM Study project with the CBA (Chemical Bond Approach) project which had been developed in the United States at about the same time. He felt that the strength of CHEM Study lay in the close relationship of experiment, principle and theory, and its repeated emphasis on careful accurate observation. However, he commented that the textbook was rather lean and much use must be made of the *Teachers' Guide*, a large volume of 785 pages.

In 1967, Miller tried the course with students enrolled in a second year course after one year of traditional teaching. He had found that even bright students had difficulty with it as a first year course but that it was much more successful as a second year course and the average students then did well. He felt the text contained very good background material.

An earlier independent assessment of the course came from England. Michael Shayer (1970) of Kings School Worcester tried the course on his first year sixth form in 1964. He compared their results on "A" level questions after one year with pupils of two previous years. Although he found a 12 per cent to 15 per cent upgrading of the results, his comments on the course are most interesting.

> Far from being pleased with these results the pupils increasingly showed signs of irritation and contempt. They did not feel they were ever really being tested and this applied equally to those of lower ability . . .
>
> After a year of this programme, I found that nearly all ability to use knowledge in unfamiliar situations, had been conditioned out of my pupils. In the context of science this should not happen . . . . Learning is many things of which producing the correct verbal response is a very small part. . . . . It is also seeing the connection between things . . . (there is a terrible feeling of being in blinkers in the CHEM Study course) (Shayer, 1970, p. 55).

In view of the CHEM Study's stated goals this was a very serious condemnation.

Again he said:

> Both the text of the chapters and the style of the questions seem based on B.F. Skinner's theory of learning which . . . says that a student has learned something when to the appropriate stimulus, he can produce the correct verbal response. It is all very reminiscent of the old primary school mathematics (Shayer, 1970, p. 55).

Shayer also criticized the attempt to cover too much ground. "Perhaps Chapter 18 shows the 'acquaintanceship' method at its worst. All areas of organic chemistry in ten days in ten lessons simply produced bewilderment" (p. 56). (British Columbia allows 15 hours at the end of the grade 11 course.)

Shayer maintained that the intentions of CHEM Study were not achieved.

> So high is the proportion of theory which the students have to take from above, that this section* contradicts, in its organization, the avowed intentions of the course to 'give a correct and non-authoritarian view of the origin of chemical principles . . . and opportunity for discovery' (Shayer, 1970, p. 56).

A gloomy note was sounded by Johnstone (1974). In an evaluation of a new Scottish chemistry syllabus, he referred to the American "concern bordering on panic" in the 1950's

---

*Last third of the book, part of Chemistry 12 in British Columbia, except Chapter 18.

which sparked the CHEM Study and the CBA projects. "These efforts", he wrote, "by their successes and even by their ultimate failure did stimulate new thinking in chemical education" (p. 21). Some of this new thinking is touched on in the next section as it relates to the choice of topics covered in the new courses and the relative difficulty of mastering certain concepts.

Concept Mastery and Difficult Topics

Shayer's experience with CHEM Study was perhaps one reason for his undertaking further research. In September 1970, he took as his starting point for a five-year course for children between the ages of 11 and 16 years, the mental capacities of this age range. He used Piaget's description of the stages of mental development as a framework in which to assess the Nuffield "O" level chemistry course. He asked whether Piaget's stages had anything to tell us about what can and should be part of a chemistry course at each age level, and what is unlikely to be grasped until the student is older. In this age span the two applicable stages are those known as the "concrete" and the "formal" operational stages. Shayer suggested that those topics which require only deduction from observations made in experiments belong properly to the stage of concrete operations. But topics requiring the formulation of a theory - for example, of chemical combination in which weight is preserved while all other properties of the elements disappear - is, he suggested, symbolic and not concrete.

Ingle and Shayer (1971) pursued this theme in more detail, analyzing individual topics from the course in terms of the minimum conceptual levels required to sustain interest and to expect good comprehension. Their conclusions should be of considerable assistance to chemistry teachers everywhere in helping them to select suitable course material for their own classes. The fact that Piaget placed the concepts of probability, correlation and proportion in the formal operational stage gives us a clue as to the difficulties with atomic theories and problems involving the multiple proportion laws, and all calculations involving equations.

Johnstone (1971), who was mentioned earlier for his views on the two experimental programmes in the United States, has approached the question of difficult topics from a different standpoint entirely. He studied the new Scottish chemistry syllabus (introduced in 1962 as a voluntary alternative) through the eyes of the pupils themselves. By the simple procedure of asking them to rate each topic on a list as either "easy to grasp", "difficult to grasp", or "never grasped", he sought to answer some of the same questions as Shayer. He also was concerned with the problem of maturity, but his studies in this area were less effective. He compared the students' ratings with their actual performance on objective tests and found good agreement between them.

Differences between the Scottish and the Nuffield courses prevent a perfect match in their findings. Johnstone places ion-electron exchange equations and redox reactions high on his difficult list. Shayer and Ingle classify the periodic table and energetics and classification of the elements as high level concepts. Both schemes place molarities, calculations from equations and proportionality problems in this category, which leads us to ask the question as to whether the difficulty is one of psychological immaturity or inadequate mathematical grounding - a topic which has been touched on in Chapter V, in discussion of the differences in student achievement between the first and second semesters.

Other writers on teaching methods (Frankel, 1959; Smeltz, 1956; Williams & Koflsche, 1967) seem to be largely concerned with improving techniques. They often anticipated the developments just discussed by their intuitive knowledge of these factors, born of experience and understanding of young minds.

VITA

Surname: GREENWOOD Given Names: LAVINIA MARY

Place of Birth: HERTFORD, ENGLAND

Date of Birth: 22 SEPTEMBER 1920

Educational Institutions Attended, with Dates of Entering and Leaving:

MIDDLESEX HOSPITAL, LONDON	1940	to	1941
UNIVERSITY OF VICTORIA	1960	to	1969
UNIVERSITY OF VICTORIA	1970	to	1979

Degrees, Diplomas, Etc., Awarded, with Dates and Names of Institutions:

MSR	1941	Society of Radiographers
CSRT	1955	Can. Society of Radiological Tech.
B.Sc.	1969	University of Victoria

Honors and Awards:

Graduate Scholarship, University of Victoria, 1974

Publications:

Actions and Reactions, A Report to the Department of Education for the Province of British Columbia on The 1972 Survey of Chemistry 11 With Special Reference to the Effect Of Semestering and Mathematical Preparation Upon Boys and Girls 1976

PARTIAL COPYRIGHT LICENSE

I hereby grant the right to lend my thesis or dissertation (the title of which is shown below) to users of the University of Victoria Library, and to make *single copies only* for such users or in response to a request from the library of any other university, or similar institution, on its behalf or for one of its users. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by me or a member of the University designated by me. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Title of Thesis/Dissertation

A STUDY OF CHEMISTRY 11 WITH SPECIAL REFERENCE TO THE EFFECTS OF THE SEMESTER SYSTEM AND SEX DIFFERENCE UPON ACHIEVEMENT.

Author [redacted]

Signature

LAVINIA MARY GREENWOOD

Name

27 april 1979

Date