

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

A COMPARISON OF THE  
MOST SUCCESSFUL AND THE  
LEAST SUCCESSFUL STUDENTS IN A  
TEACHERS COLLEGE LANGUAGE CURRICULUM COURSE

by

JOHN KENNEDY MILLAR

A thesis presented in partial fulfilment of the requirements  
for the degree of

MASTER OF ARTS IN EDUCATION

Massey University

1970

## PREFACE

This study was conducted with the co-operation of the 1969 second-year students of Palmerston North Teachers College. I am also grateful to the Principal, the staff of the College and in particular my colleagues in the Education Department, for their help and the use of College facilities. I am especially grateful to Mr A. Forrest, Senior Lecturer in Education, Palmerston North Teachers College, without whose encouragement and assistance this work may not have been completed, and to Mr D. McAlpine, Senior Lecturer in Education, Massey University, for his help in planning and executing this study.

March 1970

J.K. Millar

## C O N T E N T S

	<u>Page</u>
1. NATURE ... ..	1
2. HYPOTHESIS ... ..	4
(i) Specific Hypotheses ... ..	4
(ii) Definition of Terms ... ..	6
3. THE COURSE ... ..	8
4. SAMPLE ... ..	10
5. INSTRUMENTS ... ..	11
<u>A</u> Standardized Tests ... ..	11
(i) The Group Test of High Grade Intelligence, A.H.5 ... ..	11
(ii) The Minnesota Teacher Attitude Inventory ... ..	12
(iii) The Sixteen Personality Factor Questionnaire ... ..	15
(iv) The Self Evaluation Form for Language Teaching ... ..	18
<u>B</u> Other Data Collected ... ..	21
(i) The Course Rating ... ..	21
(ii) The College Rating ... ..	21
(iii) Practice Teaching Rating ... ..	22
<u>C</u> Fakeability ... ..	24
6. DATA ... ..	25
(i) Statistical Method ... ..	25
(ii) Age ... ..	25
(iii) Sex ... ..	27
(iv) A.H.5 ... ..	29
(v) M.T.A.I. ... ..	31

## C O N T E N T S

	<u>Page</u>
(vi) 16 P.F. Test ... ..	32
(vii) Self Evaluation Form for Language Teaching ...	35
(viii) College Rating ... ..	41
(ix) Teaching Rating ... ..	42
7. CONCLUSIONS ... ..	45
BIBLIOGRAPHY ... ..	48

### APPENDICES

A : The objectives, topics of study, teaching methods and textbooks for the Curriculum Study, "Teaching Language in the Primary School" ... ..	50
B : Statement made to students seeking their co-operation	54
C : Self Evaluation Form for Language Teaching ...	56
D : 16 P.F. Profiles ... ..	60
E : Worked Example of the Major Statistical Procedure Used	63

## *1. Nature*

Throughout the recent years of teacher training in New Zealand there appears to have been a number of popularly held beliefs, seldom recorded in written form, unsupported by evidence, but never-the-less persistent, concerning the worth of the training given, and the quality of the students who undertake it.

Some of these may be summarized as:

- (1) Women students, as a group, are superior to men in
  - (a) intelligence,
  - (b) academic success,
  - (c) teaching ability.
- (2) Older students who come to Teachers College after a period in another occupation are more successful than younger students in both academic courses and practice teaching.
- (3) There is a low, if any, positive correlation between success in academic course in the College and success in practice teaching in the schools.
- (4) Personality is the most important factor for success in teaching.

This study examines certain aspects of these beliefs through statistical analyses and evaluation. Using the academic results of a College course in language curriculum as the independent variable, the study examines,

1. Whether certain dimensions of intelligence and personality are correlated with success or failure in the samples selected.

2. Whether there is a change in self evaluation of attitudes towards teaching language and knowledge of methods between the beginning and the end of the course.
3. Whether results of the course are indicative of results in the
  - (a) College academic rating
  - (b) College practice teaching rating.

A survey of available literature failed to discover any published investigations that have objectives similar to the above, although studies by Brim (1966), Davis and Satterly (1969), Holland (1960), Li (1969) and Tarpey (1965) investigate various relationships that are relevant to this study. The most relevant is that by Li (1969) in which she investigates the connections between attitudes towards teaching, as measured by the Minnesota Teacher Attitude Inventory (M.T.A.I.) and intelligence, as measured by the A.H.5, with marks in practical teaching and the theory of education examination of two groups of students. The findings show significant relationships between the M.T.A.I. and end-of-year theory examination and also between the M.T.A.I. and practical teaching assessments. There was no significant difference in intelligence between the most successful students and the least successful in teaching practice.

The limitations in the relevance of this study appear to be that:

- (1) it concerns Chinese graduate students in the Department of Education in the University of Hong Kong so that conclusions may not be valid cross-culturally, and
- (2) the tests used were administered without translation or adaptation and the scores translated to American norms.

In addition to the above there are many studies that examine teacher

personality characteristics.

Gage (1964, p.506) says,

"The studies falling within the scope (of teacher personality and characteristics)... are too numerous for individual mention, much less for adequate treatment".

However, the literature relating to teachers in training is much less voluminous and reference will be made to selected studies at appropriate times.

## 2. Hypothesis

### GENERAL HYPOTHESIS

That there are significant differences in intelligence, personality, age and sex between students who are successful and those who fail in the Language Curriculum course taught at Palmerston North Teachers College in 1969.

### (i) SPECIFIC HYPOTHESES

- h 1.* Students who are successful have significantly higher scores on the Minnesota Teacher Attitude Inventory than those who fail.
- h 2.* Students who are successful have significantly higher scores than those who fail on both the pre and post test results of the Self Evaluation Form for Language Teaching.
- h 3.* Students who are successful have significantly higher scores in the amount of change in a positive direction in the pre and post test of the Self Evaluation Form for Language Teaching, than students who fail.
- h 4.* Students who are successful and those who fail have significantly different scores in each of the factors of the Sixteen Personality Factor Questionnaire with the exception of Factor B (intelligence).

*h 5.* Students who are successful have significantly higher scores on the Group Test of High Grade Intelligence A.H.5, than those who fail.

*h 6.* Students who are successful are older than those who fail.

*h 7.* Women students are significantly more successful than men.

*h 8.* Students who are successful have a significantly higher practice teaching rating than those who fail.

*h 9.* Students who are successful have a significantly higher College rating than those who fail.

(ii) DEFINITION OF TERMS

For the purposes of this study:

1. "College" shall be defined as the Palmerston North Teachers College during the 1969 academic year.
2. "Success" shall be defined as students who were more than one standard deviation above the mean of the total score of the Language Curriculum course results. These results comprised course work marks which received a 40% weighting, and final examination results which received a 60% weighting as shown in the table below. The course results were used to calculate the course "rating".

Course Work (40% weighting)

<u>Course Work</u>	<u>Number During Course</u>	<u>Marking Scale</u>	<u>Course Marks</u>
Essay	2	1 - 10	50
Minor Test	8	1 - 10	30
Lesson Plan	2	1 - 10	20
		<b>Total</b>	<b>100</b>

Final Examination (60% weighting)

<u>Type of Question</u>	<u>Number</u>	<u>Marks</u>
Multiple Choice Items	40	40
Short Answer	20	20
Essay	2	40
	<b>Total</b>	<b>100</b>

3. "Failure" shall be defined as students who were more than one standard deviation below the mean of the total score of the Language Curriculum Course results (defined above).
4. "Course Rating" shall be defined as the official College results of the academic courses on a 1 - 5 point scale of which 1 is the highest score and 5 is the lowest score.
5. "College Rating" shall be defined as the overall assessment on the 1 - 5 point scale, of academic work, calculated from all course ratings for each student in that year, and supplied to the Palmerston North Teachers College Council and to the New Zealand Department of Education.
6. "Practice Teaching Rating" shall be defined as the final result for practice teaching work calculated by the Principal Lecturer in Charge of Practical Training from 15 independent ratings for each student made by College Lecturers, Head-teachers and Class-teachers over the two year period.

### 3. The Course

Palmerston North Teachers College requires that all students undertake, and complete, a study of the teaching of language in the primary school, this course being known as a Curriculum Study in Language. A senior staff member has responsibility for the writing and teaching of this course which is moderated by the Head of the Education Department who also appoints a team of four lecturers to assist in the teaching.

Since all students undertake the same course, it may be considered as the independent variable, and factors that contributed to success or failure, other than the teaching methods employed by the College, may be sought.

A copy of the objectives of the course, the topics for study, and methods of teaching, obtained from College records, is supplied in Appendix "A".

#### METHODS OF COURSE EVALUATION

Since the course rating is used as the basis of the selection of the samples it is appropriate to describe the methods of evaluation at this point.

##### 1. Course Work

Course work, which received a weighting of 40% in the rating, was made up of:

- (a) Two essays - all were marked by the senior lecturer in charge of

the course.

- (b) Eight objective-type tests - set by the team as a group, administered and marked by the lecturers in charge of seminar groups.
- (c) Two lesson plans - all students were interviewed individually by one lecturer who assessed the work during the interview.

## 2. Examination Work

At the end of the course students sat a three-hour examination during which they were required to answer:

- (a) 50 multiple-choice items
- (b) 20 "short answer" questions
- (c) 2 essay question, selected from 5 options

All essay answers were marked by the senior lecturer in charge of the course and the remainder of the examination by the lecturers in the team. The compiling of the course rating was carried out by the senior lecturer in charge of second-year studies.

## *4. Sample*

On February 10th 1969 the second-year group of 244 students began a curriculum study, "Language Teaching in the Primary School". During the course seven students left the College, reducing the total number to 237 who fulfilled all course requirements and offered themselves for examination.

Following College procedures, students were allotted a course rating in which there was a weighting of 60% for the examination work and 40% for course work. Using the total score of the course results from which the course rating was obtained the mean and standard deviation were calculated. Students more than one standard deviation above the mean were selected as group one (success), and those more than one standard deviation below the mean were selected as group two (failure). Group one contained 33 students while there were 32 in group two.

### Age

The age range of the students in group one is from 18 yrs to 39 yrs (mean 22 yrs), and the range in group two is from 18 yrs to 22 yrs (mean 19 yrs)

### Sex

Of the 33 students in group one, seven are men and 26 are women.

Of the 32 students in group two, 16 are men and 16 are women.

## 5. Instruments

### A Standardized Tests

- (1) The A.H.5 Group Test of High Grade Intelligence (Heim, A.W., 1956) - administered after one week of course work.
- (2) The Minnesota Teacher Attitude Inventory (Cook, Leeds and Callis, 1951) - administered on the final day of course work.
- (3) Sixteen Personality Factor Questionnaire. "The 16 P.F. Test", Form B. (Cattell, R.B. and Eber, H.W. 1957) - administered after five weeks of course work.

### Informal Rating Scale

- (4) Self Evaluation of Knowledge in Language Teaching - administered before and after the course - see copy, Appendix "C".

### Further Data Collected

- (5) Practice Teaching Rating.
- (6) College Rating.

### THE SELECTION OF THE INSTRUMENTS

#### (i) The A.H.5 Group Test of High Grade Intelligence

In considering intelligence, Teachers College students may be regarded as a pre-selected sample from the top 50% of a cross section of the general population. Thus it is desirable to use a test that provides

greater discrimination at the higher levels than is provided by more general tests of intelligence. The author of the test, Heim (1957), claims that the "A.H.5 is a group test of general intelligence, designed for use with selected, highly intelligent subjects. It is intended for adults, such as students and research workers, and potential entrants to the university and the professions" (p.2).

Keats (1959) agrees in that "the test seems to have wide application for testing adults of high ability".

The test has a high retest reliability since, according to Keats (1959), several investigations have shown that retest reliability, with at least six months between tests, has not been less than .90.

The A.H.5 is in two parts, but Keats goes on to point out that, despite the retest reliability, four different correlations between parts one and two ranged from only .49 to .62. Thus, he concludes, "the prospective user would probably be better served by treating the two parts as separate tests".

In light of the demands on students' time, this advice was accepted and Part One only was administered after one week of the course had elapsed.

(ii) The Minnesota Teacher Attitude Inventory

Gage (1963) comments that the M.T.A.I. is "by far the most popular instrument for the measurement of teacher attitudes", and it would appear that this is equally true of the investigations of the attitudes of teachers in training.

Investigations that concern the use of the M.T.A.I. with teachers in

training appear to fall into two main divisions. The first of these examines changes in attitude associated with course work and practical training while the second is concerned with correlational studies of the M.T.A.I. and ratings received in practice teaching.

Investigating changes of attitude after six months of course work, Callis (1950) found significant changes of attitude in a "desired direction" (Gage (1963), p.509).

In correlating the M.T.A.I. and teaching ratings of teachers in training, Stein and Hardy (1957) found that six of eight correlations were significant beyond the .05 level and concluded that "student teacher attitudes are measured by the M.T.A.I. with a fair degree of validity and reliability" (p.326).

On the other hand, three studies, Sandgren and Schmidt (1956), Oelke (1956) and Fuller (1951) are critical of the M.T.A.I. because they found no significant correlation between M.T.A.I. scores and teaching ratings.

However, on the basis of the reports of these studies it seems that an alternative, and equally logical finding, could be that teaching ratings, themselves, are not reliable statistical measures. This point will be considered further in the discussion of the practice teaching rating used in this study.

Studies by Evans (1958) (1967), Day (1959), Herbert and Turnbull (1963) and Lipscomb (1956) all found significant levels of change that occurred during academic course work and teaching practice. Thus, evidence indicates that the attitudes of students, undergoing the experiences

provided by teacher training are likely to be subject to a significant degree of change. Secondly, when the stage of training is taken into account, attitudes towards teaching as measured by the M.T.A.I. can be tested with a fair degree of reliability. On the basis of these conclusions the M.T.A.I. was chosen to measure student attitudes for this study and will be correlated with the course rating since the time of administering the M.T.A.I. and the time of sitting the final examination was separated by only one week.

(iii) The Sixteen Personality Factor Questionnaire (16 P.F. Test)

"The 16 P.F. Test is an objectively-scorable test devised by basic research in psychology to give the most complete coverage of personality possible in a brief time" (Cattell, 1962) and is intended for use from age seventeen to mature adults. There are six forms of the test of which forms A and B are "most appropriate for the fully literate person, the person whose educational level is equivalent to the normal high school graduate" (Cattell, 1962).

The sixteen factors which the test is designed to assess have been labelled with terms that Cattell intends should be relatively free of connotations of "good" and "bad" (Gage, 1963). The "bipolar descriptions of source traits" (factors) provided by Cattell (1957) may be summarized thus:

Factor	Label	Description
A	Sizothymia, A- v.s. Affecto thymia, A+	Detached, critical, cool v.s. Warm-hearted, easy going, participating.
B	Less intelligent, B- v.s. More intelligent, B+	Concrete-thinking, lower scholastic mental capacity v.s. Abstract thinking, bright.
C	Lower ego strength, C- v.s. Higher ego strength, C+	Emotionally less stable, easily upset v.s. Mature, calm, faces reality.
E	Submissiveness, E- v.s. Dominance, E+	Mild, accommodating, conforming v.s. Independent, aggressive, stubborn.
F	Desurgency, F- v.s. Surgency, F+	Sober, prudent, serious v.s. Happy-go-lucky, impulsive, enthusiastic.
G	Weaker super-ego strength, G- v.s. Stronger super-ego strength G+	Expedient, evades rules, feels few obligations v.s. Conscientious, persevering.

Cont...

Factor	Label	Description
H	Threctia, H- v.s. Parmia, H+	Shy, restrained, diffident v.s. Venturesome, socially bold, spontaneous.
I	Harria, I- v.s. Promsia, I+	Tough-minded, self reliant, realistic v.s. Tender-minded, dependent, sensitive.
L	Alaxia, L- v.s. Protension, L+	Trusting, adaptable, free of jealousy v.s. Suspicious, self-opinionated.
M	Praxernia, M- v.s. Autia, M+	Practical, careful, conventional v.s. Imaginatively creative, careless of practical matters.
N	Artlessness, N- v.s. Shrewdness, N+	Forthright, natural, sentimental v.s. Shrewd, calculating, worldly.
O	Untroubled adequacy, O- v.s. Guilt proneness, O+	Confident, placid, self assured v.s. Apprehensive, depressive, troubled.
Q1	Conservatism, Q1- v.s. Radicalism, Q1+	Conservative, respects established ideas v.s. Experimenting, critical, liberal, analytical.
Q2	Group adherence, Q2 v.s. Self sufficiency, Q2+	Group dependent, a follower v.s. Prefers own decisions, resourceful.
Q3	Low integration, Q3- v.s. Self sentiment control, Q3+	Undisciplined self conflict, careless of protocol, follows own urges v.s. Controlled, socially precise, follows self image.
Q4	Low ergic tension, Q4- v.s. High ergic tension Q4+	Relaxed, tranquil, unfrustrated v.s. Tense, frustrated, overwrought.

Each of the above factors is tested by ten to thirteen items and the raw score is converted to sten scores which may be used to draw a test "profile".

Both Adcock (1959) and Gage (1963), in their reviews, regret the comparatively infrequent use of the 16 P.F. by educational research.

Although it would be reasonable to expect that frequency of use may have

increased since 1963 the 16 P.F. does not appear to be widely used, especially in the area of teacher training. Seven studies relevant to this investigation were located of which five examined the correlation between the 16 P.F. and teaching practice marks, one, the 16 P.F. and education theory exam results, and one examined the relationship with both theory and practice results. The findings of these investigations are summarized below.

Factors of the 16 P.F. Test that Appear to have a Significant Relationship with High Marks in Teaching Practice

Investigation by:	A	C	E	F	G	H	I	L	M	N	O	Q1	Q2	Q3	Q4
Davis and Satterly (1969)					*		*		*		*				*
Erikson (1954)					*				*		*			*	
Hadley (1954)				*	*					*					
Start (1966)	*		*					*	*			*		*	
Warburton <u>et al.</u> (1963)					*		*							*	
Tarpey (1965)	*				*	*			*						

Factors of the 16 P.F. Test that Appear to have a Significant Relationship with High Marks in Theory of Education Examinations

Investigation by:	A	C	E	F	G	H	I	L	M	N	O	Q1	Q2	Q3	Q4
Holland (1960)			*	*	*	*	*					*		*	
Warburton <u>et al.</u> (1963)		*			*							*			

\* Significant Relationship

In the six studies that investigate the relationship between the 16 P.F. test and high marks in teaching practice, factors G, M and Q3 were found to appear significant by three or more researchers, and factors A, I and O appeared significant in two studies. Factor Q2 was not found to appear significant in any of the studies but there was no agreement on the significance of the remaining factors.

Two studies investigated the significance of the relationship between the 16 P.F. test and high marks in education theory examinations. Both found that factors G and Q1 appeared to be significant and neither found that factors A, L, M, N, O, Q2 or Q4 appeared to be significant. There was no agreement concerning the remaining factors.

In concluding his review of the validity and reliability of the 16 P.F. test, Adcock (1959) says, "Evidently, despite the reputation of questionnaire methods as unreliable, this test does succeed". Thus, the 16 P.F. was selected as the most suitable test of personality for use in this study and Form B was administered to all students.

#### (iv) The Self Evaluation Form for Language Teaching

##### Purpose

In his introductory chapter, Anderson (1964) offers a "partial inventory of teaching knowledge" (p.4 f.f.) which he suggests may be used by students to check their progress during the course. Since Anderson was the main, though not the only text used during the language course, his "Self Evaluation Form" appeared to provide the basis of a useful method of testing both the attitudes that students held towards the teaching of language and their self-concept of knowledge of, and competence in,

teaching methods. Consequently, Anderson's inventory was changed and enlarged to meet the specific objectives of the course (see Appendix "B").

#### Content and Organization

The objectives of the course were arranged in five headings and these were placed in the approximate order of importance and emphasis given to each in the course. Fifty-five items were grouped under these headings so that the number of items under each reflected the emphases of the course in terms of lecture time spent on each topic.

<u>Headings</u>	<u>Number of Items</u>
(1) Written expression and associated skills. (spelling, handwriting, grammar)	14
(2) Instructional Reading.	13
(3) Children's Literature.	7
(4) Oral Language and associated skills - (listening, speaking)	7
(5) General	14

The heading "General" included items on language development, advice to parents, and all-inclusive items such as, "I can explain how reading, writing, listening and speaking are related" (Item 5, Appendix "B").

The items were arranged on the form in random order and students were asked to rate their present level of skill and knowledge on a

five-point scale. Forty items were intended to sample the students' self-evaluation of their knowledge and skills in teaching the various aspects of language. The remaining fifteen items were intended to sample students' attitudes towards teaching language in terms of personal satisfaction gained.

### Trial

A draft form of the inventory was tried with a group of 18 first year students and they were asked to report any difficulties they found in completing the form.

Eleven of the trial group reported difficulty with the phrase "I receive satisfaction..." (e.g. Item 48, "I receive satisfaction from the teaching of reading"). The item was revised to read "I enjoy teaching reading" and similar changes were made in the other items of the group. Minor changes were made to clarify the directions of the test and the revised form was tried with a second group of 18 first-year students. No difficulties were reported. This revised form was administered to all second-year students before and after the Language Curriculum Course.

B Other Data Collected(i) The Course Rating

At the conclusion of each academic course held within the College the Head of each Department was required to submit to the Principal a "Course Rating" on a 1 - 5 point scale which had a weighting of 60% for the final examination mark and 40% for marks gained in course work. There was an attempt to normalize the rating by directing the Heads of Departments that the five categories should be interpreted as meaning that, out of a representative group of one hundred students, the following percentages would receive each rating.

<u>Quality of Performance</u>	<u>Outstanding</u>	<u>Above Average</u>	<u>Average</u>	<u>Below Average</u>	<u>Limited</u>
Rating	1	2	3	4	5
% receiving each rating	5	20	50	20	5

The Head of Department was to keep in mind, although not necessarily adhere exactly to the above scale.

The Language Curriculum course was rated by the above procedure but to obtain a sufficient spread of marks the total score of course results was used as the basis of the selection of successful and failing students as outlined on pp.6-10.

(ii) The College Rating

The Principal took the mean of all course ratings for each student and supplied these to the New Zealand Department of Education as the "College Rating" which was intended to show the overall performance of

the student in the academic work of the College. It is not the purpose of this study to investigate the statistical validity of the procedure used to obtain the rating but, rather, to investigate the relationship between success or failure in the Language course with the overall academic performance as indicated by the College rating.

(iii) Practice Teaching Rating

During the two-year College course students were posted to primary schools for teaching practice for five periods of four weeks. On each posting each student was assessed by three people; a visiting College lecturer, the Headmaster of the school and the class teacher to whom the student was posted. At the conclusion of the posting the three independent assessments were collated by the Principal Lecturer in Charge of School Practice to produce a "Posting Rating". All assessments and the posting rating were required to be on the scale outlined in "Course Rating" above.

From the outline of assessment procedure and also from the studies of procedures in the investigations of Fuller (1951), Oelke (1956), Sandgren and Schmidt (1956) and Wiseman and Start (1965) there is no doubt that the assessment of students engaged in practice teaching is subjective in nature. It is possible that important contributing factors to success on each occasion are the relationships that develop between the student and the class teacher and the degree to which the student is prepared to conform to the requirements of the school and the college. Thus it would seem that the teaching rating, because of its subjective nature, should not be regarded as a reliable statistical measure. However, the methods of assessment of practice teaching in the College, at the time of this

study, involved fifteen independent ratings for each student with an attempt to normalize the rating scale. The fifteen ratings were reduced in six steps to a single normalized rating by the Principal Lecturer in Charge of School Practice. Again, it is not the purpose of this study to investigate the statistical validity of the rating but rather to investigate the relationship between success or failure in the Language course with the Practice Teaching Rating.

### C Fakeability

The susceptibility of self-report instruments to faking is well known, thus, since three of the instruments used in this study are of the self-report type, the problem of fakeability must be raised.

The M.T.A.I. is one instrument which has been studied for its susceptibility to faking. Callis (1950, p.727) concludes that "the Inventory was found to be only slightly susceptible to attempts to fake good", but Coleman (1954) and Stein and Hardy (1957) found significant changes in score when students were instructed to fake. In spite of these findings Stein and Hardy (1957, p.329) conclude that "this does not mean that the test is susceptible to faking, it means rather that the test is adequate in revealing a biased or prejudiced attitude towards children from either extreme position".

Sorenson (1956) investigated the effect of asking subjects to sign their names on the answer sheets of tests. He found that subjects who signed their names appeared to fake significantly less than those who did not sign their names, although both groups received the same instructions to fake. The difference in faking between the two groups was significant at the .01 level.

The conclusions drawn from the reviews of these investigations were that:

- (1) faking is possible in the three self-report instruments used in this survey but that;
- (2) this may be reduced by asking all students to sign their names on all answer sheets.

Students were also reminded before commencing each test that the results would not effect the course rating and would not be placed in College records.

## 6. Data

### (i) STATISTICAL METHOD

The method employed in this study involved the calculation of the mean of each set of data for both Group 1 and Group 2, followed by the application of a "t" test to the difference between the group means to find the level of significance.

To find the value for "t" a pooled SD was computed by the formula.

$$SD = \sqrt{\frac{\sum (x_1 - M_1)^2}{(N_1 - 1)} + \frac{\sum (x_2 - M_2)^2}{(N_2 - 1)}}$$

The standard error of the difference was then calculated by the formula:

$$SE_D = SD \sqrt{\frac{N_1 + N_2}{N_1 N_2}}$$

"t" was then obtained by dividing the actual difference between the means by the  $SE_D$  and, by entering the table of "t" for the appropriate degrees of freedom the level of significance was found. (Garrett, 1958 pp.223 - 225, formulas 57, 58; Table D, p.449) and Paradine and Rivett, 1960 pp.111 - 113). (See Appendix "E" for worked example).

### (ii) AGE

The following table summarizes the statistical analysis of age of the

students in the two groups. The groups were compared in three ways:

- (i) total group ("combined") with total group,
- (ii) men with men, and
- (iii) women with women.

Sex	Group	Age Range	M	SD	$M_1 - M_2$	"t"	P
Combined	1	18 - 38	22.09	6.66	3.06	2.55	< .02
	2	18 - 22	19.03	1.07			
Men	1	18 - 23	20.86	1.4	1.49	2.8	< .02
	2	18 - 22	19.37	1.08			
Women	1	18 - 38	22.42	7.09	3.70	2.07	< .05
	2	18 - 22	18.69	0.93			

These results indicate that Group 1 students have a higher mean age than Group 2 students, the difference being significant at between the .02 and .01 level. Thus the probability that the difference occurred by chance is less than 2 in 100.

The men in Group 1 have a higher age mean than the men in Group 2, the significance of the difference approaching the .01 level. (P .02, 2.52; P .01, 2.83; Garrett, 1958 p.449).

The women in Group 1 have a higher age mean than the women in Group 2, the significance being less than .05.

Comments:

The results tend to support the Specific Hypothesis H6, viz that the

successful students were, as a group, significantly older than the students who failed. This finding appears to have implications for the selection of students for teacher training. The Selection Committee of the College has periodically expressed reservations about accepting older students who have not undertaken academic work for several years and especially about accepting older married women with family commitments.

On the basis of this study such doubts appear to be unfounded. On the other hand, it may be argued that older students who are accepted for training are a highly selected sample and would, therefore, be expected to succeed.

(ii) SEX

Group 1 contained 7 men and 26 women while Group 2 contained 16 men and 16 women. The ratio of men to women in the total population of the second-year group was 1 : 2.9, thus a chi-square test was applied to find the significance of the difference between the observed and the expected distribution of the sexes in each group.

The formula used was:

$$\chi^2 = \sum \left[ \frac{(f_o - f_e)^2}{f_e} \right]$$

in which  $f_o$  = the number of men and women actually in each group.

$f_e$  = number of men and women expected in each group on the basis of the ratio men to women in the total year group.

(Garrett, 1958 p.253)

By entering table E (Garrett, 1958 p.450) with the computed value of

chi square and the appropriate number of degrees of freedom (in this case  $df = 2$ ), chi square was evaluated.

	Men	Women	Total
Group 1	7 (8)	26 (25)	33
Group 2	16 (8)	16 (24)	32

Calculation of  $\chi^2$

$$1^2 + 8 = .125$$

$$1^2 + 25 = .04$$

$$8^2 + 8 = 8.0$$

$$8^2 + 24 = 2.66$$

$$\chi^2 = 10.825$$

$$P < .01$$

Thus the difference in the distribution of the sexes between Group 1 and Group 2 was highly significant,  $P$  being less than .01.

To determine where this level of significance lay the chi-square test was applied to the two groups separately:

$$\begin{aligned} \text{Group 1: } \chi^2 &= .125 + .04 \\ &= .165 \end{aligned}$$

$P$  is not significant

$$\begin{aligned} \text{Group 2: } \chi^2 &= 8 + 2.66 \\ &= 10.66 \end{aligned}$$

$$P < .01$$

There is no significant difference between the observed and expected distribution of the sexes in Group 1. However, there is a high degree of

significant difference between the observed and expected distribution of the sexes in Group 2.

#### Comments

These findings only partially support hypothesis H.7 in that the ratio of men to women in Group 1 is not significantly different from the ratio of men to women in the total population of second-year students. However, there are significantly more men and fewer women in Group 2 than could be expected. The range of the rating in the Language Curriculum course is similar for both men and women but the distribution for men is positively skewed. Thus hypothesis H.7 is supported in as much that there are more men and fewer women than expected in the group who failed.

It could be useful to examine this factor further by statistically analysing the data for the total year group by applying a chi-square test to students above the median and a second to students below the median of the course rating.

#### (iv) A.H.5

The mean and standard deviation of scores on the A.H.5 Group Test of High Grade Intelligence was calculated for each group. A "t" test was applied to the difference between the means of the two groups.

The results of the analysis of the data were:

Sex	Group	M	S.D.	$M_1 - M_2$	"t"	P
Combined	1	13.76	6.18	0.23	0.17	>.1
	2	13.53	4.26			
Men	1	11.29	3.79	3.02	1.84	.1
	2	14.31	3.57			
Women	1	15.21	4.69	2.37	1.58	>.1
	2	12.75	4.72			

The data was statistically analysed in three ways; the men and women in Group 1 with the men and women in Group 2 and by sex for each group. The difference between the means for the men in the two groups was significant at the .1 level. The difference between the means for the remaining two comparisons was not significant.

#### Comments

The findings show that there is no significant difference in intelligence, as measured by the A.H.5 between the two groups; thus hypothesis H.5 is rejected. It would appear that measured intelligence was not a differentiating factor between the two groups.

It seems that further research is required into the relationship between intelligence and achievement in academic courses at the tertiary level. It may, for example, be interesting to make an item analysis of the test to find whether there is a significant difference between the course rating and the scores on items that require mathematical reasoning and items that require verbal reasoning.

(v) THE MINNESOTA TEACHER ATTITUDE INVENTORY

The means of the M.T.A.I. scores for each group were calculated and a "t" test applied to the difference between the means to determine the level of significance.

Sex	Group	M	S.D.	$M_1 - M_2$	"t"	P
Combined	1	50.63	22.4	22.96	5.03	< .01
	2	20.97	25.15			
Men	1	39.14	10.33	27.02	4.30	< .01
	2	12.12	22.78			
Women	1	53.65	23.89	23.84	3.1	< .01
	2	29.81	24.3			

The M.T.A.I. scores were thus compared in three ways. The scores for men and women in Group 1 were compared with the scores for men and women in Group 2, the scores for men in Group 1 were compared with scores for men in Group 2 and the scores for women in Group 1 were compared with scores for women in Group 2. Results for all three comparisons showed that P was less than .01. The difference between the means of the two groups is therefore highly significant.

Comment

The above results support the findings of Evans (1958) (1967), Day (1959), Herbert and Turnbull (1963) and Lipscomb (1956) who found significant relationships between academic results and the M.T.A.I. scores.

It is concluded that students who were successful in the Language Curriculum course tend to have warmer, more permissive inter-personal

relationship with pupils and be more satisfied with teaching as a vocation than students who failed, this conclusion being based on the statement of the purpose of the M.T.A.I. provided by Cook, Leeds and Callis (1951, p.3). Thus, hypothesis H.1 is supported.

(vi) THE 16 P.F. TEST

The means of the sten scores for each of the personality factors were compared in the following ways: combined group 1 with combined group 2, group 1 men with group 2 men and group 1 women with group 2 women. In addition inter-group comparisons were made between group 1 men and women and group 2 men and women. These are shown in the following table.

		16 P.F. Test Factors														
Sex	Group	A	C	E	F	G	H	I	L	M	N	O	Q1	Q2	Q3	Q4
Combined	1 c.f. 2			.1	<.01	<.02		.01		<.01			<.01		<.01	
Men	1 c.f. 2			.1	.05	0.1	<.01	<.02		<.02	.05	.05	<.01			.02
Women	1 c.f. 2				<.01	<.02		<.01		<.01			<.01		<.01	
Men c.f. Women	1	.01									.05					
Men c.f. Women	2			.1									.01		.1	

In addition to knowing the level of significance of the difference between the means it is important to note the direction of movement away from the mean range.

### Test Profiles

Cattell and Eber (1957) provide a 16 P.F. Test Profile which acts as a convenient way of finding by inspection the deviations that go beyond the mean range in either direction. Group profiles and separate profiles for men and women are provided in Appendix "D".

The profile of the groups, when men and women are combined, showed a significant difference in Factor F (Desurgency F - v.s. Surgency P+) in that Group 2 deviated beyond the mean range of the test norms towards P+. Group 2 deviated beyond the mean range of Factor I (Harris I- v.s. Premsia I+) towards I-. In Factor M (Praxernia M- v.s. Alaxia M+) Group 2 deviated beyond the mean range towards M+. Similarly Group 2 deviated beyond the mean range in Factor Q1 (Conservatism Q1- v.s. Radicalism Q1+) towards Q1+. Although group results differed significantly on Factor Q3 (Low Integration Q3- v.s. Self Sentiment Control Q3+) with Group 1 being nearer Q3- and Group 2 being nearer Q3+, neither group deviated beyond the mean range of the test norms.

When the men of the two groups were compared, the men in Group 1 deviated beyond the mean range towards Factors H+, I+, M+, N+, O-, Q1+ and Q4 minus showing that as a group they tended to be more venturesome, sensitive, imaginative, shrewd, confident, liberal and relaxed than men in Group 2.

Group 2 men deviated beyond the mean range towards factors G-, H-, I-, and Q1-, showing that as a group they tend to be more undependable, timid, realistic and conservative than the men in Group 1.

When the women of the two groups were compared the women in Group 1

deviated beyond the mean range of the test norms towards Factors M+, Q1+ and Q3+ showing that as a group they tend to be more imaginatively-creative, liberal and self controlled than women in Group 2.

Group 2 women deviated beyond the mean range towards Factors F+, G- and Q3-, showing that as a group they tend to be more impulsive, expedient and have less self discipline than the women of Group 1.

### Comments

These findings support the results of Holland (1960) and Warburton, Butcher and Forrest (1963) in that both of these studies found significant relationships between Theory of Education examinations and factors G and Q1 of the 16 P.F. This study also found factors G and Q1 to be significant. The results of this study also found with Holland (1960) that factors P, H, I and Q3 were significantly related to academic success.

Neither Holland (1960) nor Warburton, et al. (1963) found factor M to be significantly related although it was found to be significant at the .01 level in this study. Since it is "M+" that is concerned and one description of this factor is "imaginatively-creative" (Cattell and Eber, 1957), it is possible that teaching methods encouraged in New Zealand place more reliance on the teachers' imagination and initiative than is the case in Britain where teaching may be more "text-book oriented".

### Comparison of Men and Women in Each Group

When the men and women in Group 1 were compared factors A and N were significantly different for the sexes. In Group 2 men differed from women in factor Q1. It is possible that these differences are due to cultural

differences between America and New Zealand since American norms were used.

### Findings

Hypothesis H4 was not fully supported but neither can it be rejected on the basis of the evidence. Significant relationships were found for factors F, G, I, M, Q1 and Q3. No significant relationships were found for factors A, C, E, L and Q2. Factors H, N, O and Q4 were found to be significant for men.

#### (vii) SELF EVALUATION FORM FOR LANGUAGE TEACHING

##### (1) Mean Response Rating

The mean response was calculated for each of the 55 items of the inventory. The mean of these means was then obtained as the "mean response rating". This procedure was carried out with both the pre and post inventory results for each group. A "t" test was applied to the difference between the mean response ratings of both the pre and post test situation of the inventories.

	Group	M	S.D.	$M_1 - M_2$	"t"	P
Pre test	1	2.89	0.95	.25	1.47	> .1
	2	3.14	0.86			
Post test	1	1.8	0.59	.23	2.09	< .05
	2	2.03	0.56			

The difference between the mean response ratings of the two groups on the pre test was not significant. The difference between the groups on the post test was significant beyond the .05 level.

### Comment

The question of what the mean response rating shows must be considered. The mean of a set of means could seem a rather doubtful procedure to use. However, it can be argued that the purpose of the inventory was to obtain the mean ratings for the group on each of the 55 items in the inventory. Since the items were selected to reflect the objectives and the emphases of the Language course the mean of the 55 means gives an overall rating that provides a measure of the attitude of the group as a whole towards knowledge and skills in Language teaching.

The results show that at the time of the pre test there was no significant difference between the overall attitudes of the two groups but that by the time of the post test situation Group 1 rated their attitude more highly than did Group 2, the difference being significant at beyond the .05 level.

Thus hypothesis H2 is only partially supported. Successful students had significantly higher scores than students who failed on the post-test of the Self Evaluation form. However, successful students did not have significantly higher scores in the pre test.

### (2) Group Change

The mean response rating of the pre test was compared with the mean response rating of the post test for each group to determine the significance of the change in rating achieved by each group.

When comparing the mean response rating for the pre test with that for the post test, both groups achieved a change of rating that was highly significant.

Group	Inventory	M	S.D.	$M_1 - M_2$	"t"	P
1	Pre test	2.89	0.95	1.09	7.2	< .01
	Post test	1.8	0.59			
2	Pre test	3.14	0.86	1.11	7.9	< .01
	Post test	2.03	0.56			

#### Comparison of Change in Group Means

The mean change between the pre and post tests was calculated for each group and a "t" test applied to the difference between changes in group means.

Group	M	S.D.	$M_1 - M_2$	$t_{63}$	P
1	1.06	0.16	0.02	0.4	> .1
2	1.08	0.24			

Thus there was no significant difference between the two groups in the change of group mean between the pre and post tests.

#### Comment

Both groups showed a highly significant change in self rating between the pre and post test situation. Therefore it may be concluded that the students of the sample, irrespective of success or failure in the course rating, felt that the experiences provided by the course, and the associated practical work with children, had significantly increased their confidence in their knowledge and skills in Language teaching. However, when the amount of change shown by the two groups was compared, the difference between

the two was not significant. This is surprising when it is considered that there was a significant difference between the two group in post test results. There are two possible explanations.

Firstly, it is possible that the instrument failed to discriminate a real degree of change between the two groups.

Secondly, factor F (Desurgency -- Surgency) of the 16 P.F. showed a significant difference between the two groups at the .01 level. It is possible that, since Group 1 tended to be more sober and prudent than Group 2 who tended to be more enthusiastic, happy-go-lucky and impulsive, Group 2 over estimated their level of knowledge and skill. The accuracy of the Group 2 response may be of doubtful validity. Only further work with the instrument itself and a more exhaustive study of successful and failing students would determine the accuracy of these observations.

Hypothesis H3 is rejected in that students who were successful did not have significantly higher scores in the amount of change in the Self Evaluation Form.

### (3) TOPIC RATINGS

The 55 items of the inventory were regrouped under the five topic headings: Written Expression, Reading, Children's Literature, Oral Language and General (see description of instrument p. 19). The mean of each response was calculated and the mean of the means under each topic heading were taken as the "topic ratings". "t" tests were applied to the differences between each of the topic ratings of the two groups.

Gp1 c.f. Gp2	Topic Ratings				
	Written Language	Reading	Oral Language	Literature	General
Pre test	.05	> .1	> .1	> .1	> .1
Post test	< .01	> .1	< .02	< .01	> .1

On the pre test Group 1 obtained a higher rating for Written Language than Group 2, the difference being significant at the .05 level. There was no significant differences between the ratings of the two groups on the remaining four topics.

In the post test Group 1 had higher ratings for Written Language and Children's Literature, significant beyond the .01 level, and a higher rating for Oral Language, significant at the .02 level. The differences between the ratings of the two groups for Reading and General were not significant.

#### (4) SELF EVALUATION OF SATISFACTION GAINED FROM LANGUAGE TEACHING

The 15 items intended to measure the feelings of personal satisfaction gained from teaching Language (see description of instrument p.18) were used to obtain a mean rating of satisfaction for each group. A "t" test was applied to the difference between the means of the two groups.

	Group	M	S.D.	$M_1 - M_2$	$t_{15}$	P
Pre test	1	2.27	0.89	0.28	1.04	> .1
	2	2.55	0.64			
Post test	1	1.83	0.56	0.22	1.5	> .1
	2	2.05	0.58			

Thus no significant difference between the mean rating of satisfaction of the two groups was found in either the pre test or the post test.

Changes in Mean Rating of Satisfaction between the Pre and Post Tests

The mean rating of satisfaction of the pre test was compared with the mean rating of satisfaction for the post test in both groups, and the difference was tested for each group.

Group	Test	M	S.D.	$M_1 - M_2$	$t_{15}$	P
1	Pre	2.27	0.89	0.44	1.69	>.1
	Post	1.83	0.56			
2	Pre	2.55	0.64	0.5	2.27	.05
	Post	2.05	0.58			

When the means of the pre and post tests were compared Group 1 showed no significant change in the mean rating of satisfaction but Group 2 showed an increase in the mean rating of satisfaction that was significant at the .05 level.

Comment

Results show there were no significant differences between the two groups in feelings of satisfaction gained from teaching Language in either the pre or post tests. Moreover, Group 1 showed no significant change in feelings of satisfaction between the pre and post test while Group 2 showed a significant increase in satisfaction. There are three possible explanations for this:

1. The instrument failed to measure the students' feelings of satisfaction.

2. Group 1 students had already developed positive attitudes towards teaching Language due to incidental experiences on practice teaching postings prior to the Language Curriculum course. These attitudes were not significantly changed as a result of experiences provided by the course. On the other hand, Group 2 students felt more secure due to their increased level of knowledge and skill and therefore expected greater satisfaction from Language teaching experiences.
3. The significant difference in the P.F.16 desurgency - surgency factor resulted in an unreliable result for Group 2.

#### Masking

It could have been possible for significant changes to have been masked by a distribution of positive and negative changes in "satisfaction" results. In fact, this did not occur as a negative change of  $-.06$  occurred in only one of the 55 items of the inventory. Thus masking of significant change was not a possible explanation.

#### (vii) COLLEGE RATING

The distribution of the college rating for each group is shown in the following frequency table.

Rating	Frequency Group 1	Frequency Group 2
1	9	-
2	15	1
3	9	7
4	-	18
5	-	6

The statistical analysis of this data of total group scores and by sex for each group is as follows:

Sex	Group	M	S.D.	$M_1 - M_2$	"t"	P
Combined	1	2.0	0.74	1.9	5.13	< .01
	2	3.9	0.76			
Men	1	2.0	0.5	1.75	7.29	< .01
	2	3.75	0.56			
Women	1	2.0	0.78	2.06	9.36	< .01
	2	4.06	0.45			

The difference between the means of each group is highly significant ( $P < .01$ ) in each case when Group 1 men and women are compared with Group 2 men and women, Group 1 men with Group 2 men, and Group 1 women with Group 2 women.

#### Comments

These findings support hypothesis H.9 in that the difference between the mean of the two groups was highly significant. Thus, students who were successful in the Language course tended to have above average college ratings while students who failed the Language course tended to have below average college ratings.

#### (ix) TEACHING RATING

The distribution of the teaching rating for each group is shown in the following frequency table.

Rating	Frequency Group 1	Frequency Group 2
1	4	-
2	17	1
3	10	9
4	2	19
5	-	3

A statistical analysis of this data for each group, and also by sex for each group, is as follows:

Sex	Group	M	S.D.	$M_1 - M_2$	"t"	P
Combined	1	2.30	0.76	1.45	8.06	< .01
	2	3.75	0.66			
Men	1	2.0	0.53	1.69	6.04	< .01
	2	3.69	0.67			
Women	1	2.38	0.8	1.37	7.6	< .01
	2	3.75	0.65			

The difference between the means of each group is highly significant ( $P < .01$ ) for each comparison made.

#### Comment

These findings support hypothesis H.8 in that students who were successful in the Language course tended to have higher teaching ratings than students who failed. It would seem that the popular belief that students who succeed in college academic work are not successful in the classroom may be unfounded. However, it should be noted that only 4

students from the total of 12 who received a teaching rating of 1 were in Group 1 and only 3 of the total of 9 students who received a teaching rating of 5 were in Group 2. Although the hypothesis is supported it does not follow that there is exact match between top academic students and students rated 1 on practice.

## 7. Conclusions

The conclusions, based on the statistical analysis of the data and related to each of the specific hypotheses, may be summarized as follows:

*h1.*

The hypothesis that students who are successful in a language curriculum course have significantly higher scores on the M.T.A.I. than those who fail was supported.

*h2*

That students who are successful in a language curriculum course have significantly higher scores than those who fail on both the pre and post test results of the Self Evaluation Form for Language teaching tended to be true of only the post test situation.

*h3*

That students who are successful in a language curriculum course have significantly higher scores in the amount of change in a positive direction in the pre and post test of the Self Evaluation Form for Language Teaching, than students who fail was rejected on the basis of the evidence.

*h4*

That students who are successful in a language curriculum course and those who fail have significantly different scores in each of the factors of the 16 P.F. was found to be true for six of the factors, namely F, G, I, M, Q1 and Q3.

*h5*

That students who are successful in a language curriculum course have significantly higher scores on the Group Test of High Grade Intelligence A.H.5 than those who fail was rejected.

*h6*

That students who are successful are older than those who fail was supported.

h7.

That women students are significantly more successful than men was partially supported in that there were more men in the group which failed than expected on the basis of the ratio of men to women of the total sample.

h8.

That students who are successful have a significantly higher practice teaching rating than those who fail was supported.

h9

That students who are successful have a significantly higher College Rating than those who fail was supported.

Most of the hypotheses of this study were supported although this could be expected since the two groups of students were selected on the basis of scores in the Language Curriculum course that were more than one standard deviation above or below the mean. Thus, the samples tended to be highly selected and some significant differences between the groups could be expected with reasonable confidence.

In addition to the methods of statistical analysis used, other procedures could have been applied. For example, the application of a multiple regression equation could have been useful to determine the "weight" of each of the variables as possible contributing factors to success (Garrett 1958, p.410). Further, the inter-relationship of the factors in the study meant that it is possible that some significant factors were partialled out.

There is, at present, a lack of a body of systematic research into the training of teachers in New Zealand and further research in this area would be of value. The data for this study suggests interesting possibilities for further work. One such area is the relationship of intelligence to achievement in tertiary levels of education. This study found the difference in intelligence scores between successful and

unsuccessful students, as measured by the A.H.5, was not significant, thus supporting the findings of Li (1969), but the question remains whether the level of measured intelligence is significant as a factor of successful achievement in other aspects of Teachers College work. Why is it that students with a higher measured intelligence do not tend to also have higher levels of achievement?

Further work on the personality factors of New Zealand teacher trainees could also prove profitable. The weight of opinion and research is against the use of personality measures as an aspect of trainee selection, yet there is some agreement between the findings of this study and others in both America and Britain. It is possible that further research may show consistent relationships between certain personality factors and success in teacher training that may have implications for the methods of selection of trainees. Further, the significant differences noted in this study between the sexes on certain personality factors raises the question of the suitability of American norms to the New Zealand situation. Overall, the investigation of relationships between attainment levels and the factors of personality, attitudes and intelligence of New Zealand students is, as yet, in an early stage of development.

BIBLIOGRAPHY

- Adcock, C.J. (1959) Review of the 16 P.F. In *Buros* (1959) 5 : 112.
- Anderson, P.S. (1964) Language Skills in Elementary Education. New York : MacMillan.
- Arvidson, G.L. (1966) Learning to Spell. Wgtn : N.Z.C.E.R.
- Auckland Education Board (1963) Reading in the Primary School. Auckland : The Board.
- Brim, B.J. (1966) Attitude Change in Teacher Education Students. J. Educ. Res. 59, 10 pp.441-445.
- Buros, O.K. (Ed.) (1959) The Fifth Mental Measurements Yearbook. New Jersey : Gryphon.
- Callis, R. (1950) Change in teacher-pupil attitudes related to training and experience. Educ. Psychol. Measmt. 10 pp.718-727 In Gage (1963).
- Cattell, R.B. and Eber, W.H. (1957) Handbook for the Sixteen Personality Factor Questionnaire "The 16 P.F. Test". Illinois : Institute for Personality and Ability Testing.
- Cattell, R.B. and Eber, H.W. (1962) Manual for Forms A and B : Sixteen Personality Factor Questionnaire. Illinois : Institute for Personality and Ability Testing.
- Coleman, W. (1954) Susceptibility of the Minnesota Teacher Attitude Inventory to "Faking" with Experienced Teachers. Educ. Admin. Superv. 40 : pp.234-237. In Gage (1963).
- Cook, W.W., Leeds, C.H. and Callis, R. (1951) The Minnesota Teacher Attitude Inventory. New York : Psychol. Corp.
- Davis, T.N. and Satterly, D.J. (1969) Personality Profiles of Student Teachers. Brit. J. Educ. Psychol. 39, pp.183-7.
- Day, H.P. (1959) Attitude Changes of Beginning Teachers after Initial Teaching Experience. J. Teacher Educ. 10, pp.326-328.
- Erikson, H.E. (1954) A Factorial Study of Teaching. J. Exp. Educ. 23, pp.1-39. In Gage (1963).
- Evans, K.M. (1958) An Examination of the Minnesota Teacher Attitude Inventory. Brit. J. Educ. Psychol. 28, pp.253-257.
- Evans, K.M. (1967) Teacher Training Courses and Students' Personal Qualities. J. Educ. Res. 10, pp.72-77.
- Fuller, E.M. (1951) The Use of Teacher-Pupil Attitudes, Self-Rating and Measures of General Ability in the Pre-service Selection of Nursery School-Kindergarten-Primary Teachers. J. Educ. Res. 44, pp.675-686. In Gage (1963)
- Gage, N.L. (Ed.) (1963) Handbook of Research in Teaching. Chicago : Rand McNally and Co.
- Garrett, H.E. (1958) Statistics in Psychology and Education. London : Longmans.

- Hadley, S.T. (1954) A Study of the Predictive Value of Several Variables to Student Teacher Success as Measured by Student Teaching Marks. Teach. Coll. Bull. State Teachers Coll. Indiana. 60 : pp.1-10. In Gage (1963)
- Hall, V.C. (1965) Former Student Evaluation as a Criterion for Teacher Success. Jnl. Exp. Educ. 34, 1, pp.1-19.
- Heim, A.W. (1956) Manual for the Group Test of High Grade Intelligence A.H.5. Slough : National Foundation for Educ. Research in England and Wales.
- Herbert, N and Turnbull, G.H. (1963) Personality Factors and Effective Progress in Teaching. Educ. Rev. 16, pp.24-31. In Evans (1967)
- Holland, J.L. (1960) The Prediction of College Grades from Personality and Aptitude Variables. Jnl. Educ. Psych. 51, pp.245-254.
- Keats, J.A. (1959) Review of the Group Test of High Grade Intelligence A.H.5. In Buros (1959) 5 : 390.
- Li, A.K. (1969) Student Attitudes and Teacher Training Performance. J. Educ. Res. 12, pp.60-3.
- Lipscomb, E.E. (1966) A Study of the Attitudes of Student Teachers in Elementary Education. J. Educ. Res. 60, pp.159-163.
- N.Z. Department of Education (1963) Primary School Syllabus : Language in the Primary School : English. Wgtn : Govt. Printer.
- N.Z. Department of Education (1965) Primary School Syllabus : Reading. Wgtn : Govt. Printer.
- N.Z. Department of Education (1966) Suggestions for Teaching English in the Primary School. Wgtn : Govt. Printer.
- Oelke, M.C. (1956) A Study of Student Teacher's Attitudes Towards Children. J. Educ. Psychol. 47, pp.193-196. In Gage (1963).
- Paradine, C.G. and Rivett, B.H.P. (1960, 2nd Ed.) Statistical Methods for Technologists. London : English Universities Press.
- Sandgren, D.L. and Schmidt, L.G. (1956) Does Practice Teaching Change Attitudes Towards Teaching? J. Educ. Res. 49, pp.673-680. In Gage (1963).
- Sorenson, A.G. (1956) A Note on the "Fakability" of the M.T.A.I. J. Appl. Psychol. 40, pp.192-194. In Gage (1963)
- Start, K.B. (1966) The Relationship of Teaching Ability to Measures of Personality. Brit. J. Educ. Psychol. 36, pp.158-165.
- Stein, H.L. and Hardy, J.A. (1957) A Validation Study of the M.T.A.I. in Manitoba. J. Educ. Res. 50, pp.321-338. In Gage (1963)
- Tarpey, M.S. (1965) Personality Factors in Teacher Trainee Selection. Brit. J. Educ. Psychol. 35, pp.140-149.
- Warburton, F.W; Butcher, H.J. and Forrest, G.M. (1963) Predicting Student Performance in a University Department of Education. Brit. J. Educ. Psychol. 33, pp.68-79.
- Wiseman, S. and Start, K.B. (1965) A Follow-up of Teachers Five Years After Completing Their Training. Brit. J. Educ. Psychol. 35, pp.342-360.

A P P E N D I X    "A"

A copy of    A    Objectives

B    Topics for Study

C    Teaching Methods

D    Textbooks

for the curriculum study

"Teaching Language in the Primary School"

---

TEACHING LANGUAGE IN THE PRIMARY SCHOOL

A. OBJECTIVES

While all objectives are interpreted in practice in behavioural terms, more generally it is the aim of the course that students:

- (1) Gain an understanding of the nature, function and development of language.
- (2) Develop an understanding of current methods of teaching language in the New Zealand primary school.
- (3) Study recent developments in the teaching of language.
- (4) Be given opportunity to practice teaching language skills in the classroom.

B. TOPICS OF STUDY

More specifically course objectives are interpreted through the following topics of study:

- (1) The nature and function of language - language as communication.
- (2) Language development from birth to adolescence - revision and extension<sup>of</sup> work in the year one Child Development course.
- (3) Language in the Primary School - a study of the requirements of the syllabus "Language in the Primary School, English".
- (4) Reading -
  - (i) Theoretical considerations.
  - (ii) Introduction to methods of instruction.
  - (iii) Current methods in New Zealand schools.

- (a) The New Zealand Scheme for reading in infant classes -  
The Language - Experience approach.
  - (b) Basic readers - method of use, strengths and limitations.
  - (c) The "Auckland" Reading Plan.
  - (d) Individualized reading in Intermediate Schools.
  - (e) S.R.A. Reading Laboratories.
- (iv) Innovations in reading instruction.
- (a) The initial teaching alphabet.
  - (b) Reading in colour.
  - (c) The O.K. Moore Typewriter Procedure.
- (5) Listening -
- (i) The nature of listening.
  - (ii) Development of listening skills.
  - (iii) Class programmes in New Zealand schools.
- (6) Oral language -
- (i) Freeing children for effective speaking.
  - (ii) Teaching speech skills (a) Infant classes.  
(b) Standard classes.
- (7) Written language -
- (i) Teaching written expression (a) Expository writing.  
(b) Creative writing.
  - (ii) The skills of written language (a) Spelling.  
(b) Handwriting.
- (8) Remedial teaching in the language programme -
- (i) Language problems of the disadvantaged child.

- (ii) Specific remedial techniques - diagnosis and treatment.
- (9) Language and subject integration -
  - (i) The integrated timetable.
  - (ii) Language in other subject areas.

#### C. TEACHING METHODS

This course will use the following teaching methods:

- (1) Mass lecture - 120 students per session.  
Time - 1 hr., Frequency - Two per week.
- (2) Seminar - 15-18 students per session.  
Time - 1 hr., Frequency - Two per week.
- (3) Demonstration by normal school staff  
- 65 students per session.  
- Frequency of use to suit topics.
- (4) Supervised teaching practice.

#### D. TEXTBOOKS

Anderson, P.S. (1964) - Language Skills in Elementary Education -  
New York : Macmillan.

Arvidson, G.L. (1966) - Learning to Spell - Wgtn : N.R.C.E.R.

Auckland Education Board (1963) - Reading in the Primary School -  
Auckland : The Board.

N.Z. Dept. of Education (1966) - Suggestions for Teaching English  
in the Primary School - Wgtn : Govt. Printer.

Cont...

N.Z. Dept. of Education (1963) - Primary School Syllabus :  
Language in the Primary School : English -

Wgtn : Govt. Printer.

N.Z. Dept. of Education (1965) - Primary School Syllabus : Reading -

Wgtn : Govt. Printer.

A P P E N D I X "B"

A copy of the statement made by the author  
of this study to all students seeking their  
co-operation in the testing programme.

---

We are all very familiar with the concept of evaluation and testing in as much that all students undergo the same course of study and at the end of it sit an examination. The examination results and the marks from tests and essays presented during the course are then turned into a course rating and this rating will vary from one to five.

Why do students show this variation in results? An obvious answer is, of course, that it depends on the amount of work that each student puts into the course work and the time spent "swotting" for the exam. But are there any other factors that contribute to success in any particular course? Has intelligence level got anything to do with it? Do the attitudes that you have towards the job of teaching as a whole and towards the course in particular have any effect on results? Do the women usually do better than the men? Is there any connection between the results of College courses and the teaching practice rating? We can all make guesses, but no-one really knows for sure. I would like to conduct an investigation that tries to answer some of these questions, but to do so means giving you some tests to assess intelligence and attitudes towards teaching. I have selected five tests that I think would be useful and I want to ask you today if you would assist me by doing these tests. There is not time to do these during committed time so they would have to be done during "off-seminar" time. It would mean giving up five study periods spaced throughout the course which is not a great deal when taken over the year.

Let me assure you that the results of the tests will be kept confidential

and that they will not affect your course or teaching ratings in any way. They will NOT be placed in the College records.

I will give you a few minutes now to discuss among yourselves whether you are happy to take the time to do these tests and then I will answer any questions.

Let me stress that if you are prepared to do the tests it is vital that everyone turns up to do each test on each occasion.

We will have a short break now for discussion.

A P P E N D I X "C"

Self Evaluation Form for Language Teaching

---

SELF EVALUATION FORM FOR LANGUAGE TEACHING

NAME.....SECTION.....SEMINAR.....

AGE.....YRS.....MTHS (1st February 1969)

Rate your present skill or knowledge for each of the following aspects of teaching language by placing a tick in the most appropriate box of the five point scale.

+ means you 'feel completely confident', or 'definitely yes', or 'very sure', or 'completely agree'.

- means you 'feel completely inadequate', or 'definitely no', or 'very unsure', or 'completely disagree'.

Example: I can use the Arvidson system of teaching spelling.

+					-
---	--	--	--	--	---

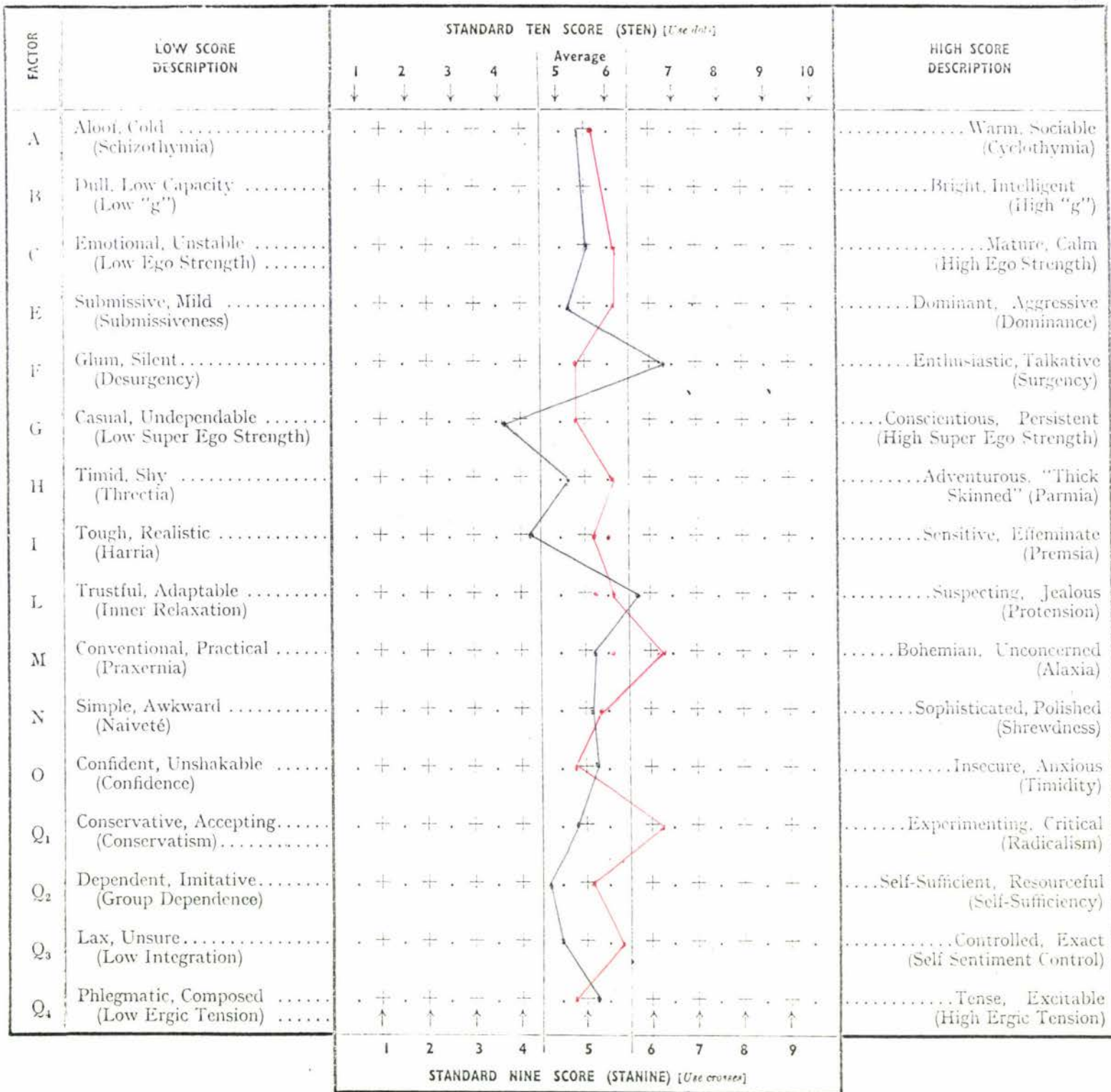
(On the form supplied to students all items were as per the example)

1. I understand the historical background of the language I teach.
2. I know how children master language before starting school.
3. I consider knowledge of how children master language before starting school to be relevant for Primary teachers.
4. I know how language instruction fits into the school day.
5. I can explain how reading, writing, listening, and speaking are related.
6. I can give a parent a reference that will help guide the recreational reading of her children.
7. I can intelligently purchase books for a classroom library.
8. I can hold children's attention as I tell a story.
9. I enjoy reading aloud to children.

10. I know how to share a picture book with little children.
11. I know how to encourage children to read good literature.
12. I know the sources of poetry that are appropriate to children at the various levels of the Primary school.
13. I can read poetry well.
14. I can recognize different types of speech defects.
15. I know how to distinguish the speech defects that I can help and those that need a speech therapist.
16. My own voice is free from irritation to those who listen to it.
17. I know how to listen to children.
18. I can discuss the importance of teaching listening skills in the classroom.
19. I can write neatly and legibly on the blackboard.
20. I can write a note to a parent and not be embarrassed by my handwriting.
21. I know the reasons why I teach handwriting the way I do.
22. I know how to help a left-handed child to write well.
23. I know why the Arvidson spelling system is organized the way it is.
24. I know how to make the spelling period interesting to children.
25. I am able to help children to write creatively.
26. I know the relationship between oral and written expression.
27. I can discuss the difficulties involved in marking children's written expression.
28. I know what type of work in written expression to expect of children in the different ability levels.
29. I understand the fundamentals of basic grammar.
30. I understand the conflict between teaching grammar and functional usage to children.

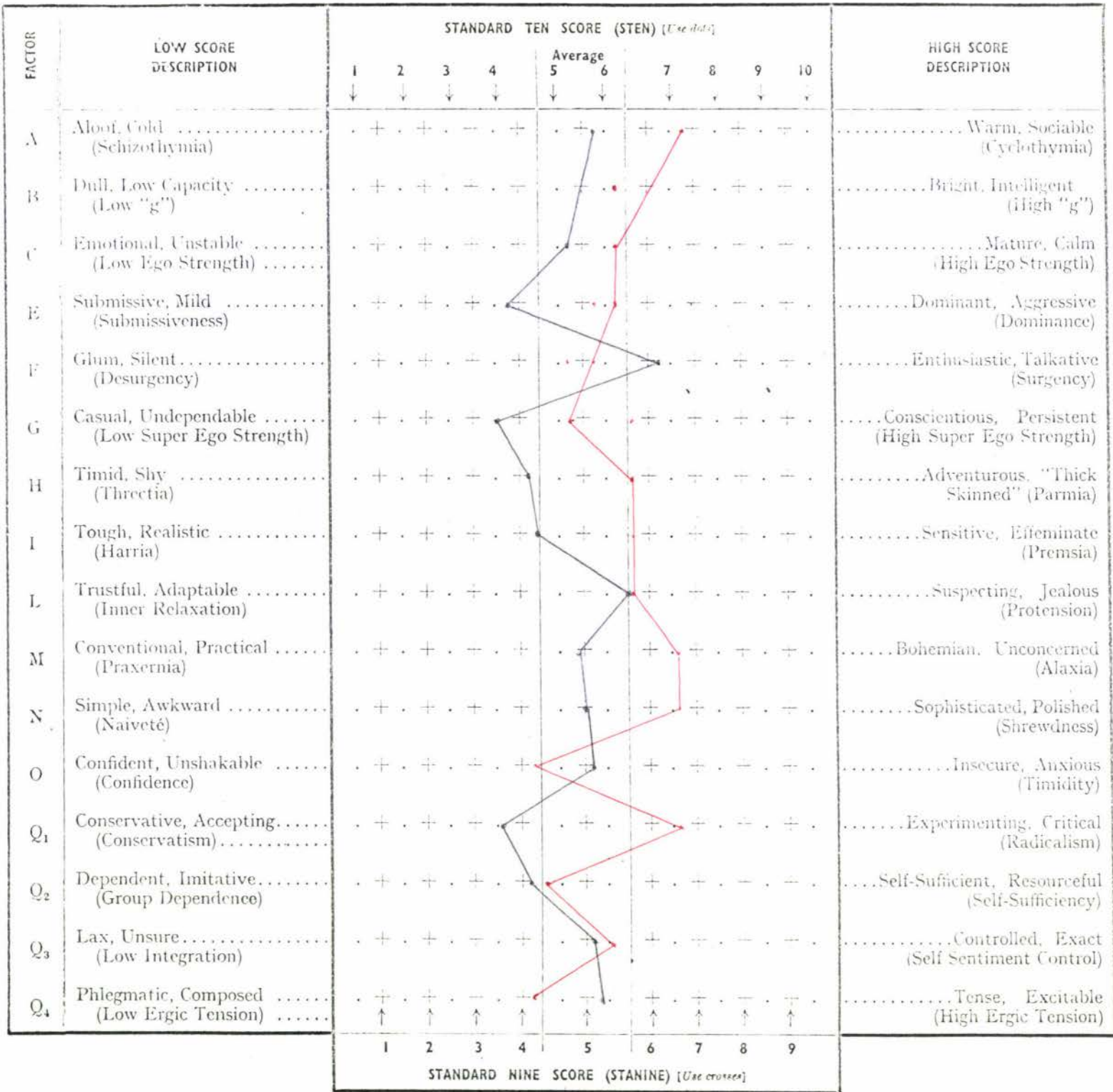
31. I understand the significance of readiness in all learning experience.
32. I can teach a beginner to read the material in the books designed for him.
33. I understand the strength and weaknesses of phonics as applied to spelling and reading.
34. I know of material designed to be used with children who are below grade level in reading.
35. I know at least one of these magazines concerning the language arts: Elementary English, The Reading Teacher, Elementary School Journal, The Instructor.
36. I enjoy helping children who have special difficulty in learning to read.
37. I can give a talk at the P.T.A. explaining the significance of the skills of language arts for children in the space age.
38. I am aware of the importance of research in education in the field of Language Teaching.
39. I enjoy teaching creative writing.
40. I expect all children to be able to read after spending two years at school.
41. I expect a child who talks well to read well.
42. I understand why it is desirable for children to learn to read and write at the same time.
43. I consider that teacher-chosen centres of interest are best for young children.
44. I can initiate the 'Auckland Reading Plan' in standard classes.
45. I can discuss the advantages of 'Basic Readers' in the teaching of reading.
46. I can describe to a parent the place of supplementary readers in the New Zealand reading scheme.
47. I can use an 'Informal Reading Inventory'.
48. I enjoy teaching reading.

49. I can recommend three picture dictionaries for use in infant classes.
  50. I can chose three tests that would be helpful when diagnosing a remedial reader.
  51. I know how to make effective use of the Country Library Service.
  52. I know how to develop purposeful discussion in an oral language programme.
  53. I believe that discussion has an important place in the language programme.
  54. I can accurately judge whether the level of difficulty of a passage of prose suits the reading level of a particular child.
  55. I understand the use of the S.R.A. Reading Laboratory.
-



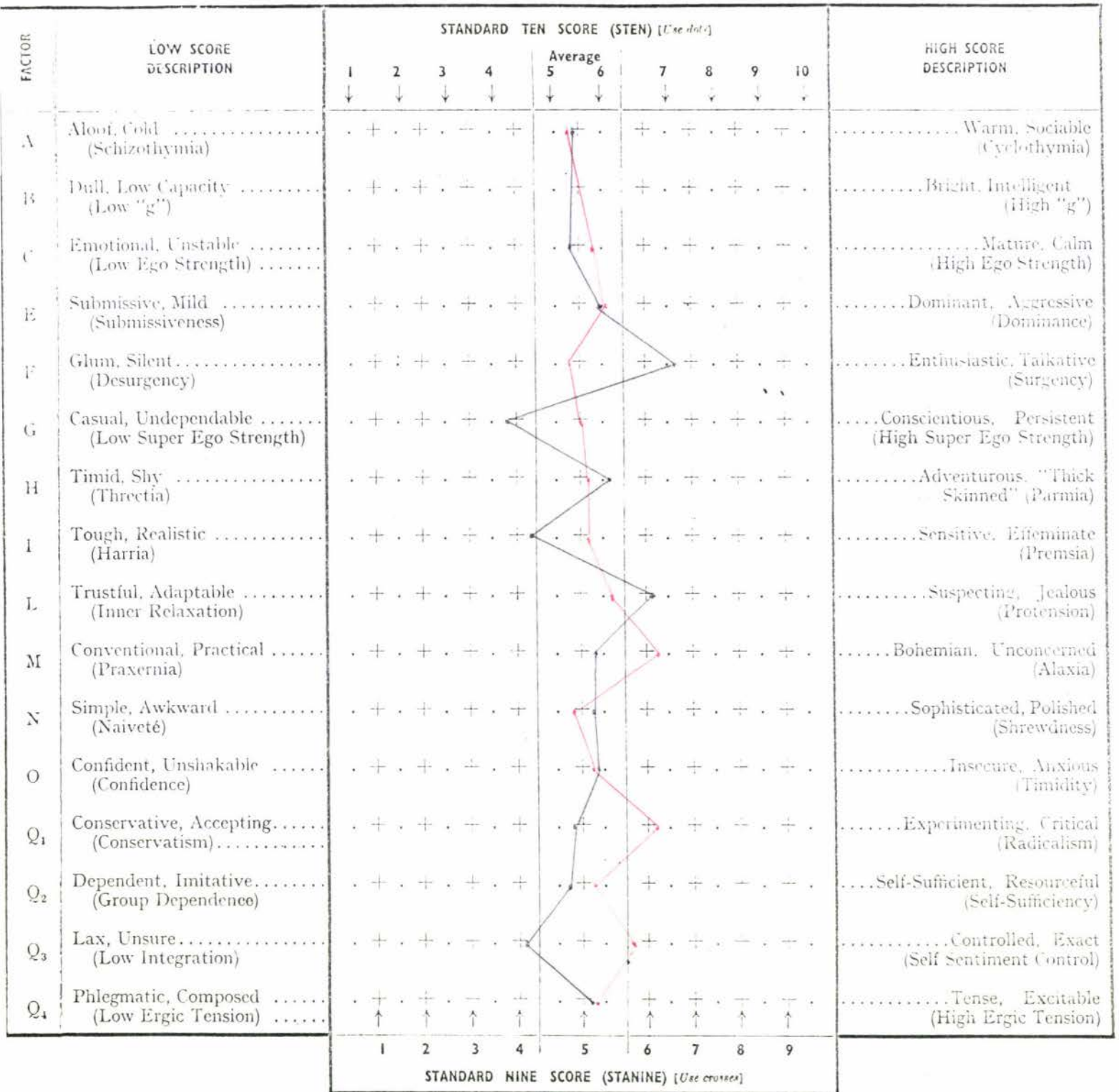
Men and Women Combined.

— Group 1.  
 — Group 2.



Men.

— Group 1.  
 — Group 2.



Women

— Group 1  
 — Group 2

Worked example: The significance of the difference between  
the means of the two groups for M.T.A.I. scores.

Group	Mean	S.D.	N	df.
1	50.63	22.4	33	32
2	20.97	25.15	32	31

$$\text{"Pooled" SD} = \sqrt{\frac{(22.4)^2 \times 32 + (25.15)^2 \times 31}{32 + 31}}$$

$$= 23.79$$

$$SE_D = 23.79 \sqrt{\frac{33 + 32}{33 \times 32}}$$

$$= 5.95$$

$$t_{63} = \frac{29.66}{5.9}$$

$$= 5.03$$

$$P < .01$$

A P P E N D I X "D"

16 P.F. Test Profiles

- (1) Men and Women Combined, Group 1 and Group 2.
  - (2) Men, Group 1 and Group 2.
  - (3) Women, Group 1 and Group 2.
-

A P P E N D I X "E"

Worked example

of the major statistical procedure used

---