THE APPLICATION OF A SUBJECTIVE PROBABILITY TECHNIQUE TO DIAGNOSTIC TESTS IN HIGH SCHOOL SUBJECT MATTER

A Thesis

Presented to

the Faculty of Graduate Studies

University of Manitoba

In Partial Fulfillment

of the Requirements for the Degree

Master of Education

by

Frederick Miller Olsen

1970

THESIS ABSTRACT

PROBLEM

The main purpose of this study was to examine the feasibility of using a subjective probability approach to testing in high school subjects. In the process, an attempt was made to estimate the relative merits of using the technique for (a) standardized tests, the results of which would be used as measures of aptitude (b) subject tests, the results of which could be used for predictive purposes and (c) diagnostic tests, the results of which would be used to expedite the review and reteaching of subject matter.

METHOD

The application of the subjective probability technique to scoring test items was assessed in four ways:

(1) The results of standardized test items used in the subjective probability method were correlated with those of the same items treated in the conventional manner by the same students.

(2) Specifically selected diagnostic test items were treated in both the conventional and the subjective probability manner by all students. These two sets of scores were also correlated.

(3) The results of both conventional and subjective probability treatments of the diagnostic tests were correlated with mid-year examination results in the appropriate subjects.

(4) Item analyses of each of the subject tests were performed to indicate the diagnostic information available from such a scoring technique.

It is generally the practice to postulate the "null hypothesis" and then, as a consequence of acquired information, to accept it, or to reject it, and accept an alternate hypothesis.

The following hypotheses were postulated:

(1) The Spearman rank order correlation between the results of conventional and subjective probability treatments does not differ significantly from zero.

(2) The Pearson product-moment correlation between conventionally treated subject tests and mid-term examination results does not differ significantly from the Pearson product-moment correlation between subjective probability treated subject tests and mid-term examination results.

The rejection of the null hypotheses would indicate that extra information might be made available from both ordinary standard test items and from items which had specifically been prepared. The subjective probability scoring technique could provide this information without destroying that available from the conventional results.

CONCLUSIONS

1. The subjective probability technique does not alter the rank of an individual from what it would be by using the conventional scoring method. This fact was true for both standardized tests and teacher-prepared diagnostic tests.

2. The subjective probability technique does not yield results which are of greater predictive value for essay-type, mid-term examinations than do the conventionally scored tests.

3. Partial knowledge can be readily assessed by means of analyses of frequency distributions of scores when the test items are scored by subjective probability techniques. This application could result in more efficient diagnostic and remedial work without altering any of the information concerning grading or prediction which is ordinarily available from multiple-choice tests.

ACKNOWLEDGEMENTS

This thesis was prepared under the direction of Dr. Peter A. Taylor, Associate Professor of Education at the University of Manitoba. He provided guidance and constructive criticism throughout the planning and writing of this study.

Thanks are also due Messrs. M. Cowie, G. Laidlaw and H. T. McCracken, who provided test items for the various diagnostic instruments and to Frank G. Garvin, Jr. who proof-read the text.

The author gratefully acknowledges these contributions.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

STATEMENT OF THE PROBLEM AND DEFINITION OF TERMS

I. INTRODUCTION

For the past several decades there has been an increasing emphasis on the use of various types of discrete-answer objective tests. Some of the advantages of such tests when compared with essay-response tests are considered to be:

(a) A wide sampling of skills and knowledge is possible;

(b) Scoring is convenient and fast per unit;

(c) Many adaptations of type of question are available;

(d) High validity, reliability and usability are possible for most purposes.

The conventional arguments in favour of the essay-response have included:

(a) The student learns to plan and express his argument coherently in essays;

(b) It takes less time and effort to prepare an essay-response test;

(c) For purposes of testing the specific techniques of writing, superior validity is achieved by the marking of the actual work.

The arguments in favour of the essay-response test may be construed as criticisms of the objective testing technique. To the list of weaknesses in the objective-test program several others could be added. The conventional multiple-choice test tends to be designed, by the insertion of specifically planned distractors, to discover what the student does not know rather than to examine and diagnose what he has learned and understands. It is as though one could infer the store of information, ability and knowledge at hand by subtracting the unknown fraction from the sum total of knowledge available.

In many areas of study and discussion it is of importance that the students realize and accept that there are no "true" answers or "untrue" interpretations. When scored in the conventional way, the multiple-choice test item tends to eliminate this pedagogical or philosophical approach.

The chance of appreciably and consistently affecting a total score by random guessing has been conclusively refuted by R. F. Graesser (1966) as statistically insignificant. However, the fact remains that when one considers individual items of a test for analysis of responses, the probability of a lucky guess is possibly quite significant. One cannot be certain, when considering a "correct" response, whether it was:

(a) a reflection of accurate knowledge of the information being examined;

(b) the resultant of a combination of some knowledge and a variety of subconscious clues (actually an infinite range of possible combinations); or

(c) merely a "lucky" guess reflecting no knowledge of the material in question.

Similarly, when an answer is one of those judged "incorrect", one cannot definitely infer that the response is due to a lack of knowledge or information. It may be that the "distractor" was selected because of a more complete knowledge of, or more perceptive thought about, the subject and a rejection of the trivial or superficial choice provided as the "true" answer.

To be sure, this latter type of item would be culled from a standardized battery because of the negative correlation which would be observed in detailed item analysis. With less formally constructed test items, incorrect conclusions may often result because of invalid assumptions. In order to discover information concerning the actual reason for a given response, personal interviews and discussion would have to be employed.

Referring specifically to those subject areas which lend themselves best to objective testing, it appears that it would be possible to extract additional information from the results of such tests without the associated disadvantages of the subjective items or the impractical expenditure of time involved in personal interviews. This procedure could be followed if details, concerning the form of and reason for the selection, were to be given by the testee. An indication of confidence in, or agreement with, any or all of the possible selections would also have to be indicated by the subject as he answers the item. While the subjective probability technique basically produces a measure of the degree of confidence in a given topic, for practical purposes of testing, one could infer partial knowledge from the results. It is necessary that this process be both simple and objective for the student to make the indication and for the tester to assess it.

This greater measure of information could of course be put to use in more effective teaching or reteaching of subject areas. If it is to be a grading device, more accurate ranking or grouping could result because of the availability of a more complete measure of the knowledge of the student. If the test were to be designed for purposes of establishing entrance standards and hence ultimately used as a predictor, it should be possible to produce more reliable regression equations or prediction tables from the results because of the more accurate and complete picture afforded by the extra information. The degree of validity of such devices should correspondingly be increased.

II. PURPOSE OF THE STUDY

Because of the lack of diagnostic information available from the results of conventionally scored objective tests, it was decided to conduct an empirical investigation into the feasibility of the subjective probability approach to testing in high school. The technique was tested both with standardized test items and with items which had been prepared with a view to using the results as diagnostic tools for purposes of achieving identified aims in specific courses.

The application of the subjective probability method of marking test items was assessed in five ways:

1. The results of the items used from standardized tests which were selected for subjective probability treatment were correlated with those same items treated in the conventional method.
2. The results of the selected items of standardized tests treated in the subjective probability way were correlated with total scores in standardized tests.
3. The specifically prepared diagnostic test items were treated in both the conventional and the subjective probability manner and the results correlated.
4. The results of both conventional treatments and subjective probability treatments of the specially prepared subject tests were correlated with the mid-year examination results in the specified subjects.
5. Item analysis of each of the specially prepared test items was performed in order to indicate the extra diagnostic information available.

III. DEFINITION OF TERMS

Discrete Answer Test Item

This is an objective multiple-choice question. No provision for explanation or extenuation is provided.

Conventional Scoring Technique

This term refers specifically to discrete answer items in which only one response is judged of value for scoring purposes. The suggestion that partial knowledge may be present in the form of guessed answers has stimulated testers to correct for guessing in order to arrive at a "true" score. The formula used for the items in this study was:

$$\text{Score} = \text{number of correct responses} - \left(\frac{w}{n - 1}\right)$$

w = number of wrong responses

n = number of choices available per item

Students are informed that there will be a correction factor applied to nullify the effects of random guessing.

Subjective Probability Scoring Technique

This term also refers to multiple-choice questions but involves the selection of one or more of the provided alternatives as an answer. The subject is then expected to associate a numerical weight with each of the alternatives selected. The weights are to be in proportion to his confidence in the appropriateness of the choice and in this study were limited to a sum total of five points per item. For practical purposes, this indication of degree of confidence has been equated with partial knowledge of the subject matter. Thus a weighting of five points on the one correct response receives maximum value. A distribution of the five points among several alternatives receives partial value for the item, while the indication of five points on an inappropriate choice receives zero credit. The formula used is the de Finetti equation which is outlined in the review of the literature.

IV. HYPOTHESES

The following null hypotheses are presented:

1. The Spearman rank order correlation between the results of conventional and subjective probability treatment does not differ significantly from zero.
2. The Pearson product-moment correlation between conventionally treated subject tests and mid-term examination results does not differ significantly from the Pearson product-moment correlation between subjective probability treated subject tests and mid-term examination results.

Because it has been the practice in classroom situations to consider T-scores, or rank in class, as a measure of a student's ability or knowledge, it was decided to use the Spearman rank order correlation as a preliminary test of the effect of the two scoring techniques. The rejection of the first hypothesis would lead to an examination of how significant the relationships between sets of scores are.

The results of progress tests are used to predict achievement. Often this is merely an extension, by means of a fixed rank order in a class, to predict final examination standings, but sometimes simple or multiple regression equations are developed for specific institutions. The rejection of the second hypothesis would lead one to determine which of the techniques produced results which were most closely correlated with the results of the essay and problem-type mid-term examinations.

V. SUMMARY OF CHAPTER I

The purpose of this study was to collect and examine empirical evidence concerning a method of extracting additional information from the results of objective tests --both conventionally-scored standardized tests and those specifically prepared to exhibit the characteristics in question. The resulting data were correlated with regular examination results with a view to indicating the value of the application of this data to prediction purposes.

A further aim of the study was to indicate the application of the subjective probability scoring technique to diagnostic purposes in specific subject areas.

VI. OUTLINE OF THE PRESENTATION

The next chapter presents a review of the literature on the related fields of:

(a) Objective test use;

(b) Results of and corrections for guessing on objective tests;

(c) The meaning of subjective probability;

(d) The application of subjective probability scoring to indicate students' degrees of confidence by their weighting of answers to test items.

Chapter III outlines the design of the investigation and relates information on the population, instruments and procedure.

Chapter IV is an analysis of the data collected and includes both item analyses of the instruments and group data correlations.

The conclusions and implications of the study are presented in the final chapter.

CHAPTER II

A REVIEW OF THE LITERATURE

In order to assess the various testing techniques, particularly to indicate the subjective probability application, an examination was made of professional publications which included: The Mathematics Teacher, Journal of School Science and Mathematics, Journal of Experimental Education, Scientific American, British Journal of Psychology, Educational and Psychological Measurement, Journal of Educational Psychology, Psychometrika and Acta Psychologica. It was decided to make a four-staged approach to the topic:

1. An introduction of types of objective testing;
2. An examination of the question of guessing;
3. The meaning generally taken for "subjective probability;"
4. A review of what has been done and suggested with the subjective probability approach.

Much has been written and spoken concerning the relative merit of various types of test items. Adams (1954) has summarized the advantages of the various types of test items. Her conclusions indicate that for most purposes the objective tests are more appropriate than the essay or supply-type items.

The effect of guessing is a significant problem in the practice and theory of testing. Shuford and Massengill (1966) claim that the conventional multiple-choice test items cannot be assessed in any way to precisely indicate the extent to which guessing has affected any given set of records. Gritten and Johan (1941) estimated the level of confidence of a subject by calculating the product moment correlation between the results of tests written under contrasting administrative directions. One set of directions indicated that the subject was not to guess, but to omit any item if he were unsure of the correct answer. The other set of directions stated that all items were to be answered. Those subjects with a large discrepancy in scores were assumed to have a high level of confidence in their own knowledge. Using the results of sixty-five items presented twice to each of 100 subjects, they calculated a positive correlation between the extent of guessing and the confidence level of the individual. Another personal characteristic which may be related to the tendency to guess on tests in the felt need for achievement.
Van der Meer (1967) calculated a negative correlation between reliability of response and the need for achievement.

It is of course possible to introduce correction factors partially to negate the effect of guessing. The results of tests which had been treated with such correction factors used in conjunction with information about the amount of experience the subject has had with similar tests and the degree of self confidence of the individual, would increase the probability of predicting the true score. By means of a formal model involving a large number of test items and a population of subjects who would eventually take all items, Shuford and Massengill (1967) postulated a mathematical relation between the score, the number of items taken, the proportion of items comprehended and the subjects' guesses, with a fixed probability of success, for the remaining items.

Shuford and Massengill'sformula indicates that the probability of success on a test is in favour of those subjects with a large "probability of success" for each item. They show (1967) however that if data are not corrected for guessing, the resulting scores represent a function of the total number of guesses. Because of these conflicting effects, Shuford and Massengill (1967) developed tables to show that even though it was not possible to show the extent of the effect of guessing on scores, if test results were to be used for predictive purposes there was merit in applying a correction for guessing formula. Without a correction factor the scores were less reliable for predictive purposes than was the assumed mean of the group being considered.

The recognition and application of theorems concerning mutually exclusive and independent events, the degree to which given events fit the mathematical model, the definition of "equally likely" events and the application of permutations and combinations to the definition are all unintentionally part of subjective probability judgments. As the term itself suggests, however, this particular aspect of probability defies objective laws in final analysis and must be treated according to the definitions accepted by the subject using it.

Ferguson (1959) says that probability may be used subjectively to refer to an attitude of doubt regarding future events. It is an assertion of a degree of confidence with regard to a future occurrence. It is evident that "probability" and especially "subjective probability" have acquired many shades of meaning not all of which are clearly distinguishable from each other. In ordinary discussion the context can generally make the meaning of the term clear, but when any word is used in a technical sense, it is important finally to arrive at a clear, unambiguous meaning for the specific application. The connection between the abstract mathematical stipulations of probability and the pragmatic content of the statement involving the term should be established.

Van Naersen (1962) considers subjective theory of probability logical to the extent that if a person has a degree of belief, p, in a state, s, then he should have a degree of belief (1-p) in the denial of s. One could say that any degree of belief is acceptable for any given statement, but then there are restrictions of the degrees of belief concerning related statements. Kyburk and Smokler (1966) consider subjective probability to be quasi-logical in that, although there is a numerical relationship between degrees of belief in related statements, the values are not unequally determined. Except in the case where one statement is a clear denial of the other, the sum of probabilities does not necessarily add up to unity.

Cohen (1957) considers all actions to be based on our private assessments of the laws of chance. This assessment depends on the past experiences to which we relate the situation presently at hand, but we all develop some subjective concepts of probability at any early age. Without relating in detail or belabouring the point of coherence, Cohen indicates that there is experimental evidence that, with increasing experience, uncertain situations are structured for closer and closer proximity with the objectivity of mathematical expectation and therefore would tend toward a maximum value of unity when mutually exclusive items are considered.

There seems little doubt that a working understanding of the laws of probability as applied to degrees of belief or confidence is developed in children much younger than high school age. Piaget (1961) says an understanding of the laws are not conclusively developed at the ages of five to seven, but Yost (1962) outlines specific experiments which show that nursery school children apply the rudiments of probability. Leake (1965) found that the elementary concepts of probability were well understood at junior high school age without any formal training. With the evidence on hand it seems reasonable to assume that high school students would have sufficient grasp of the subject to enable its efficient application in answering objective test items.

In attempts to overcome inherent deficiencies of subjective test items, many types of multiple-choice test items have been employed. Coombs (1953) suggested that subjects should be instructed to eliminate by crossing out all the alternatives which they consider wrong. An individual who knows the right answer will cross out all the wrong alternatives. One not knowing the right answer may have sufficient partial knowledge to eliminate one or more of the wrong alternatives. Such an individual should be instructed to cross out those alternatives he believes to be wrong, but not to guess among the remaining ones. In effect, he should exhibit just that knowledge of which he is sure. It is assumed, though not tested or proven, that elimination of some of the wrong choices implies partial knowledge of the answer.

The scoring procedure would be such that the odds are against guessing the correct answer. Each wrong alternative crossed out is worth a unit of credit. If the right answer is crossed out, it is given a weight of 1-k in which k is the total number of alternative. It is evident that a 5-alternative item would yield a nine-point scale ranging from + 4 to - 4.

The labour involved in scoring such items would be increased, and some time would be required to explain to students what the effects of guessing would be. There are also points in favour of such a method. No correction for guessing is required as each student sets up his own levels of assurance, and the examination of item responses could be put to diagnostic use.

Coombs, Milholland and Womer (1956) assessed the above suggested response method for items of a multiple-choice test. They found clear evidence that such scoring increased the reliability of the tests to a degree equivalent to a 20 per cent increase in the tests' effective length. They found that the experimental method discriminated among individuals in the test scores as well as by the conventional method and concluded that the increased reliability justified the longer scoring time required. The subjects being tested responded that they considered this method preferable to the usual multiple-choice marking scheme because they believed their standing here would be more consistent with their knowledge or skill. Other variations of the items in the multiple-choice test include the negative-response selection in which only the one alternative which is not true is indicated.

Ranking of alternatives in order of preference and assigning numerical indications of the degree of confidence in each selection was suggested by de Finetti (1965). He developed a method for scoring such items:

$$S = r_h + \frac{1}{2}\ (1 - \sum r^2)$$

S = item score

r_h = number of points assigned by the respondent to the response judged by item-writer to be most appropriate

$\sum r^2$ = sum of the squares of points assigned to each option.

Shuford and Massengill (1966) developed a similar mathematical equation to deal with the scores from individual items when subjective probability weighting or selecting had been applied --i.e., when the student had indicated his degree of confidence in each choice by assigning a relevant proportion of five points to each:

$$S = \frac{r_h}{(\sum r_i^{\,2})^{\frac{1}{2}}}$$

S = item score

r_h = number of points assigned by respondent to the most appropriate response.

$\sum r_i^{\,2}$ = sum of squares of points assigned to each option.

The de Finetti formula, because of its binomial form, has some value associated with responses which have indicated zero confidence in the accepted choice, except in the case when the 5/0 combination is used. The Shuford-Massengill formula, because of its monomial form, must have a zero numerator, and therefore zero score, for any combination which indicates zero confidence in the most appropriate response. In most cases the de Finetti formula assigns slightly larger values to the various degrees of confidence combinations. Jenkins (1967) provided a table indicating the values assigned to partial knowledge for each of the possible combination of weights.

Taylor and Jenkins (1969) outline a study of such an application with grade four students using S.C.A.T. tests. They indicate that the subjective probability approach is feasible with grade four students, but that practice sessions are required to familiarize the subjects with the method. They also state that little additional information was available from total scores, but that item analyses and frequency distributions of responses to item alternatives provided information which would enable educators to pin-point specific weaknesses, misinformation or lack of information and thus enable selective reteaching.

The suggestion is made that even more useful information would be forthcoming from instruments which had been designed specifically for diagnostic purposes and with the intent of using the subjective probability technique for identifying and obviating deficiencies of the students learning.

SUMMARY

The various types of conventionally treated objective tests seem to share most of the advantages of wide sampling, objective assessment and rapid scoring per item. They also have in common the deficiencies caused by the presence of undetectable guessing. While the effect on the total score can be reduced when the results are to be used for ranking or predictive purposes, a completely new approach is required if confidence is to be placed in individual item results.

The laws of mathematical probability are precise and rigourous, but when probability is treated subjectively, considerable latitude can be introduced. The application of subjective probability to indicate a degree of confidence in a judgment or choice seems to offer a useful tool for scoring objective test items. If the examiner can infer a partial knowledge related to the indicated partial confidence, he would have a means of assessing individual test items for diagnostic purposes.

CHAPTER III

OUTLINE OF THE PROCEDURE

I. INTRODUCTION

The general acceptance of the multiple-choice type of test item for objective measurement does not imply that it is optimal for all purposes. Used in its conventional form, it leaves much to be desired if a measure of partial knowledge is required. For purposes of diagnostic testing, the subjective probability method of scoring infers additional information about the subjects' knowledge, without significantly altering the rank of scores.

The students used in this study were selected from 100 high-school boys whose intelligence quotients ranged from 110 to 140. They were enrolled in university entrance courses--grades ten through twelve. The groups selected were complete classes in each of the academic areas being tested and thus they varied in number depending on the size of the class. The academic areas used in this study, chemistry, geography and physics, were selected for reasons of the convenience of the writer, as he had taught them and was therefore familiar with testing done there.

II. PRELIMINARY ASSUMPTIONS

1. The scoring method as well as all modes of responding was known by all subjects. Furthermore the subjects understood the implications of the weighting with reference to degree of confidence or certainty.
2. Subjects were interested in maximizing total scores. (In this context it was assumed that the imaginary reward suggested in the directions served as a means of explaining procedure rather than as a stimulus to succeed.)
3. The subjects understood the necessary correspondence between their own confidence and the numerical probabilities into which these must be translated.

III. TEST SUPPLIES

The following materials were used:

1. Differential Aptitude Tests of Verbal Reasoning and Abstract Reasoning (referred to in future as D.A.T.--V.R. and D.A.T.--A.R.) with answer sheets and manual.
2. Otis Quick-Scoring Mental Ability Tests. Gamma Form A.M. (referred to in future as Otis) and manual.
3. Teacher constructed objective tests in:
   (a) Grade 10 physics and geography
   (b) Grade 11 geography and chemistry
   (c) Grade 12 physics and chemistry
4. Mimeographed directions for the explanation of

subjective probability scoring.

5. Mimeographed answer sheets for the teacher-constructed subject tests.
6. Pencils and scratch paper.

IV. SPECIFIC DIRECTIONS FOR ADMINISTRATION OF TESTS

When used in the standardized way, the D.A.T. and Otis tests were written under the conditions outlined in the manuals. For all tests treated in the subjective probability manner, the prepared directions were read and the examples discussed until all questioners had been satisfied. The directions stipulated that each student read all of the five alternatives for each item and select the single best answer. Next the student was asked to indicate the measure of his confidence in the choice by putting a weight of from one to five on it. He was to distribute the remainder, if any, of 5 points among the other alternatives in such a way as to reflect his confidence in each of them. It was pointed out that the scoring formula had one factor related to the amount of knowledge (the weight on the correct choice) and another factor related to his confidence in that knowledge. If a subject had no confidence in any one alternative of a given item, his best "bet" was to place equal weights on all alternatives and thus receive 3 of 5 possible points per item. If the subject were entirely confident of the answer, he would place all 5 points on it. If the choice included the answer, he would receive the 5 points for that item. If, on the other hand, full confidence had been misplaced on a wrong alternative, he would receive zero. Any combination in the distribution was permissible, and the best score would result from choices which represented both knowledge and confidence. While both knowledge and confidence are required to produce the optimum score, partial confidence would produce middle scores and misplaced confidence, the minimum. For practical purposes of diagnostic testing, the inference is made that partial confidence indicates partial knowledge and misplaced confidence, misinformation or "negative" knowledge.

While the de Finetti formula was developed to scale the five point distribution back to a unit score maximum, it was decided to scale these scores back to a maximum of five for purposes of this study. This procedure removed some fractions from totals, and in the case of the selected items of standardized tests it produced scores with the same maximums as those of the total tests.

The de Finetti formula as indicated in the review of the literature is $S = r_h \quad \frac{1}{2}(1 - \sum r_i^2)$. If a student places maximum weight on the correct response, his score is calculated as follows:

$$S = 5/5 + \frac{1}{2}\left(1 - \frac{5^2}{5}\right)$$

$$S = 1$$

which is scaled up to 1 x 5 = 5 points. If a student placed equal weight on each of the five alternatives, his score was calculated as follows:

$$S = \frac{1}{5} + \frac{1}{2}\left(1 - 5\left(\frac{1}{5}\right)^2\right)$$

$$= \frac{1}{5} + \frac{1}{2}\left(1 - \frac{1}{5}\right)$$

$$= \frac{3}{5}$$

which is scaled up to $\frac{3}{5} \times 5 = 3$ points.

The table of item scores for the various combinations of point distribution is included in Appendix C.

V. THE TESTING

The Standardized Tests

The D.A.T. and Otis tests served two purposes. They were used as practice sessions before writing the diagnostic tests to familiarize the subjects with the subjective probability method of scoring. The results of the standardized tests when scored by the two different techniques, subjective probability and conventional, were available for comparison with the results of the diagnostic tests which had been scored in the same ways.

All grade ten subjects scored the D.A.T. tests by both methods. One half scored first by subjective probability technique and later under conventional directions, while the other half scored them in reverse order. The grades eleven and twelve subjects all had written the Otis previously and so merely repeated the tests under subjective probability rules. In all cases every question was answered under conventional directions; however, only every fifth item was scored according to the subjective probability technique.

The Diagnostic Tests

One of the purposes of this study was to prepare test items especially designed to make use of the subject's degree of confidence in his selection of an answer. Another purpose was to demonstrate the use of such items for diagnostic purposes.

Items were constructed in the various subject areas which were to be tested. The distractors were generated by selecting "wrong" answers which had regularly occurred in conjunction with "correct" responses on class tests or discussions. The particular facet being tested by each item was also noted for use during item analysis.

VI. COLLECTION OF DATA

The grade ten class was used as an experimental control group to discover any differences in related scores which could be attributed to the difference in the order of writing the D.A.T. tests under the two sets of rules--subjective probability and standardized. The A group had the D.A.T. tests (VR and AR) administered first in the standard manner, and the next day, in the subjective probability way, using five marks allocated to each of every fifth question. On the same day, the second day of testing, the B group wrote the D.A.T. tests by the subjective probability method using every fifth question as had the A group. On the third day the B group wrote both parts of the D.A.T. tests under standard conditions and rules.

The rules for subjective probability scoring were read, explained and discussed, using the examples supplied on the front pages of D.A.T. and Otis booklets. The standard rules were used when standardized tests were written under conventional conditions. In all cases, conventional and subjective probability scoring, students were allowed to set the test aside and work on another assignment when satisfied with their answers.

Every student in grade ten was asked to answer the following questions after the last section of the D.A.T. had been completed:

1. Which form of test do you prefer--ordinary checking of one answer or distributing weights according to your confidence in the choice?
2. If you could arbitrate the type of scoring to be used on your mid-term examinations which of the following would you select?
   (a) all subjective probability items
   (b) some subjective probability items
   (c) no items of this type.

Grades eleven and twelve were administered the Otis in the conventional way. During the experimental period they were allowed forty minutes to mark every fifth question by the subjective probability method and allowed to set the test aside if completed ahead of time.

In all cases it was emphasized that results would be discussed with individuals but that because of the danger of invalidating these tests for future use in the school particular items from the standardized tests would not be analysed.

Each of the subject tests was given with directions to indicate the first choice for each item. This selection was to be indicated by a check in the appropriate space in the top row of blanks on the answer sheet, appropriately numbered for that item. The students were asked to then indicate their degree of confidence in that choice by a number ranging from 1 to 5 in the second row of blanks below the check for that particular item. Finally they were to distribute the remainder of the 5 points per item in accordance with their agreement or confidence in the rest of the choices for each item. The students were reminded of the directions previously outlined for subjective probability scoring of the standardized tests, and the rules were read through quickly.

One 40 minute period was allowed for each test with the student setting the test aside and proceeding with another assignment when satisfied with his answers.

VII. TREATMENT OF DATA

Scores

The raw scores for all standardized tests were treated in the manner prescribed in the manual in order to arrive at comparisons with national norm groups--i.e., percentile ranks for D.A.T. results and intelligence quotients for the Otis results.

"Systematic" raw scores were derived by considering only every fifth question and multiplying that total by five. These were then treated as raw scores to arrive at percentile ranks and intelligence scores.

The subjective probability scores were prepared by employing the de Finetti formula and a third set of percentile ranks and intelligence scores were developed.

The subject tests, scored under the conventional technique, were marked and recorded. The scores were multiplied by five in order to have scores directly comparable with the subjective probability results. The subjective probability scores for these tests were also prepared by using the de Finetti formula.

All raw scores for all tests were ranked and T-scores for each calculated so that different test results for each student would be directly comparable without referring back to the number of scores.

Correlations

(a) Spearman's rank order correlation of raw scores was calculated between the conventional and the subjective probability scores in each test and the significance indicated. This coefficient was also compared for both grade ten groups to indicate whether the order of writing by the two methods affected the correlation.

(b) The Pearson product moment correlation was calculated to relate the subjects' results on the mid-term examinations with the scores on the diagnostic tests when treated by conventional and by subjective probability technique respectively.

Individual Item Tabulation

The teacher-constructed diagnostic tests of ten students were selected at random from each subject group and tabulated in terms of the types of combination of weights used for each item by each subject. From these data, suggestions concerning general deficiencies in class knowledge were made possible, and areas of necessary reteaching were noted. From each of the groups of ten, a single test was randomly selected and an examination of each response made with interpretation of indicated weaknesses. Possible remedial action was suggested for each of these latter cases.

VIII. SUMMARY

The students used for this study were boys of above average ability. The standardized tests used all have five alternatives per item, and the teacher-constructed diagnostic test items had the same format. One of the choices was the correct answer in each case; the other four alternatives had been selected to indicate shades of misconception.

All tests were scored by both the conventional method and by the subjective probability technique in order that student ranks achieved by the two techniques could be compared. The formula used in calculating the subjective probability scores was developed for items with only one correct answer, but the weighting of distractors provided information which assisted in diagnoses of misconceptions. Tests with items whose alternatives have gradations of "truth" might even reinforce this process, but a different type of scoring formula would have to be developed.

The scores of the tests were used in three ways:

(a) positions in classes were compared under the two techniques

(b) correlations were calculated between the sets of scores for each test and between each of the subject tests and the mid-term examination results

(c) frequency distributions of class responses to the individual items and individuals' responses to the various items were calculated.

It was recognized that if ranks between scores on the tests written by the two techniques were highly correlated, the correlation between each of these sets and the mid-term examination scores would also be highly correlated. The converse is not necessarily true, however, and for this reason both correlations were planned and executed.

CHAPTER IV

ANALYSIS OF DATA

I. CORRELATION OF TEST DATA

A summary of the standardized test results for grade ten DAT-VR and DAT-AR are included in Tables I and II. The scores reported in Table I are those achieved by the students who responded to the DAT-VR first in the conventional manner and later in the subjective probability manner and the DAT-AR in the reverse order. In Table II they are those achieved by the students who responded to the DAT-VR first in the subjective probability manner and later in the conventional manner and the DAT-AR in the reverse order. The columns are listed in the order the tests were scored. The column labelled "Score" is made up of the raw scores based on the conventional marking procedure, namely, total numbers of correct responses. The column labelled "S.P." is made up of the scores on every fifth item of the test which were scored according to the subjective probability formula. The columns labelled "Systematic" are derived from the conventional scores on every fifth item on each test--i.e., the same items that were used in the subjective probability scoring technique. In order to have a common maximum and also simplify fractions, these last two sets of scores, "S.P." and "Systematic", were multiplied by five to produce the values shown in the tables. The reason for including these total scores was to show how the "S.P." and "Systematic" scores related to the totals which would have ordinarily been used with the standardizing tables--i.e., the National Percentiles, I.Q.'s and the T-scores.

TABLE I

RAW SCORES DAT-VR, CONVENTIONAL THEN SUBJECTIVE PROBABILITY, DAT-AR IN REVERSE ORDER

	DAT-VR			DAT-AR		
Student	Score	Systematic	S.P.	S.P.	Systematic	Score
1	46	40	44	49	45	44
2	45	45	44	46	45	41
3	43	40	36	46	45	45
4	43	45	45	46	45	42
5	42	45	40	40	35	43
6	42	45	41	50	45	46
7	41	40	39	49	45	47
8	40	45	47	42	45	42
9	39	35	30	39	40	40
10	38	40	40	48	45	44
11	37	35	17	48	40	41
12	37	40	41	45	50	44
13	35	35	31	27	45	42
14	35	35	34	39	35	43
15	35	45	45	48	45	45
16	32	40	30	41	45	44
17	31	40	33	40	35	39
18	30	15	29	35	35	35

TABLE II

RAW SCORES DAT-VR, SUBJECTIVE PROBABILITY THEN CONVENTIONAL, DAT-AR IN REVERSE ORDER

	DAT-VR			DAT-AR		
Student	S.P.	Systematic	Score	Score	Systematic	S.P.
1	42	45	45	42	50	46
2	38	40	42	46	45	46
3	47	45	42	32	25	34
4	45	40	41	45	45	45
5	45	40	40	45	45	50
6	29	45	39	39	25	39
7	46	45	39	38	30	32
8	45	45	39	36	35	45
9	36	30	37	39	35	39
10	48	40	36	40	45	46
11	38	40	35	39	35	45
12	41	30	32	45	45	49
13	27	25	29	20	15	18
14	39	35	26	35	35	22
15	36	35	23	41	45	43
16	30	20	23	39	40	44
17	25	30	23	36	40	43
18	28	15	19	40	35	42

The first task undertaken was to determine whether the order of responding to and scoring the items, conventional then subjective probability or subjective probability then conventional, have any statistically significant effect on the values of recorded scores. In order to determine this effect, the Spearman rank order correlation coefficients were computed for each of the six sets of scores and recorded in Table III.

The maximum number of students in any class tested was 18 and with this number of pairs of scores being considered, the Spearman coefficient must be larger than .399 to be significantly different from zero at the 5% level of confidence or more than .564 to reach 1% level of confidence.

TABLE III

SPEARMAN'S CORRELATION TO IDENTIFY EFFECTS OF ORDER OF SCORING

Test	Order	r_s
DAT-VR	Conventional - S.P.	.589
DAT-VR	Systematic - S.P.	.627
DAT-AR	Conventional - S.P.	.806
DAT-AR	Systematic - S.P.	.830
DAT-VR	S.P. - Conventional	.627
DAT-VR	S.P. - Systematic	.669
DAT-AR	S.P. - Conventional	.652
DAT-AR	S.P. - Systematic	.554

As indicated in Table III, in all cases the correlations were significantly greater than zero at the 5% level of confidence, and, in all but the one case, at the 1% level. The Fisher transformation was used to test for significance of the difference in the correlations, and in all cases it was possible to show no significant difference at the 1% level of confidence. This observation was taken as evidence that the order of scoring did not alter the ranks and the T-scores, and percentile ranks were computed for the complete grade in Tables IV and V.

TABLE IV

D.A.T. T-SCORES

	VERBAL			ABSTRACT		
Student	Total	Systematic	S.P.	Total	Systematic	S.P.
1	72	50	55	56	56	63
2	66	60	55	49	56	54
3	66	60	55	51	70	54
4	62	50	47	60	56	54
5	62	60	59	51	56	54
6	58	50	48	66	56	54
7	58	60	51	55	43	45
8	58	60	53	66	56	70
9	58	60	66	33	33	38
10	55	50	59	60	56	50
11	55	50	50	72	56	63
12	54	50	59	60	56	70
13	54	60	66	51	56	47
14	52	45	42	47	49	42
15	52	60	36	45	43	42
16	52	60	63	40	35	37
17	52	60	59	37	43	50
18	50	50	51	56	56	59
19	49	45	28	49	49	59
20	49	40	47	45	43	42
21	49	50	53	56	70	50
22	48	50	72	47	56	54
23	45	45	44	51	56	35
24	45	45	33	55	43	42
25	45	60	59	60	56	59
26	45	50	48	45	43	50
27	43	40	53	60	56	63
28	43	50	41	56	56	46
29	42	50	45	45	43	45
30	41	30	36	55	43	40
31	40	37	33	28	28	28
32	38	45	50	35	43	33
33	35	45	47	49	56	48
34	35	35	42	45	49	50
35	35	40	28	37	49	48
36	28	30	35	47	43	47

TABLE V

D.A.T. PERCENTILE RANKS

Student	Total	Systematic	S.P.	Total	Systematic	S.P.
1	99	99	95	95	99	97
2	99	99	99	80	97	97
3	99	97	99	85	97	99
4	99	90	95	97	97	97
5	99	99	99	85	97	97
6	97	95	95	97	97	97
7	97	97	99	90	75	50
8	97	97	99	97	99	97
9	97	99	99	35	45	20
10	97	99	95	97	97	97
11	97	95	95	99	97	99
12	95	99	95	99	99	97
13	95	99	99	97	99	97
14	95	75	90	85	85	97
15	95	70	99	75	70	75
16	95	99	99	70	70	50
17	95	99	99	65	35	30
18	95	95	95	55	97	50
19	90	30	90	95	99	97
20	90	90	75	80	99	75
21	90	97	95	70	70	50
22	90	99	95	95	97	99
23	90	80	90	75	97	97
24	90	85	90	85	25	97
25	90	99	99	90	70	50
26	90	95	95	97	99	97
27	80	97	75	70	97	50
28	80	75	95	97	99	97
29	80	83	95	70	75	50
30	75	70	25	90	50	50
31	70	65	60	15	15	10
32	60	95	90	50	15	50
33	55	90	90	80	90	97
34	55	75	40	70	95	75
35	55	60	75	55	90	75
36	35	70	25	75	85	50

A summary of the results on the grade 11 and 12 I.Q. tests are included in Tables VI and VII. Like the grade 10 standardized tests, raw scores were computed in the conventional manner, and I.Q.'s calculated from the Tables for Deriving I.Q.'s on Otis Quick Scoring Mental Ability Tests (Harcourt Brace & World, Inc.). Also, as in the evaulation of the grade 10 standardized tests, a raw score of systematically selected items and a Subjective Probability raw score were produced for each student, and corresponding "I.Q.'s" calculated. In each table the first column gives the I.Q. arrived at under conventional rules and scoring. The second columns give the I.Q. derived from the results of the subjective probability scoring using the de Finetti formula applied to every fifth item. The third column, labelled Systematic I.Q., gives the I.Q. derived from the sampling of every fifth of the conventionally scored test items. In each of these last two cases, the raw scores achieved were multiplied by five, and these scores used with the conventional tables to calculate the I.Q.'s.

TABLE VI

GRADE XI I.Q. SCORES

Student	I.Q.	S.P. I.Q.	Systematic I.Q.
1	134	141	137
2	121	141	141
3	115	139	141
4	109	137	134
5	116	136	134
6	127	135	131
7	125	134	140
8	127	133	130
9	129	133	129
10	110	132	130
11	119	130	126
12	112	125	130
13	106	124	130
14	104	124	125
15	115	123	119
16	110	120	120
17	104	119	119
18	108	118	115
19	108	118	115
20	103	113	113
21	101	111	114
22	103	109	105
23	106	101	103
24	108	107	108
25	98	98	98

TABLE VII

GRADE XII I.Q. SCORES

Student	I.Q.	S.P. I.Q.	Systematic I.Q.
1	127	140	140
2	137	139	139
3	112	139	139
4	129	137	134
5	126	135	136
6	122	135	134
7	112	134	133
8	120	130	134
9	123	129	130
10	113	129	129
11	103	128	126
12	108	116	113
13	104	112	109
14	126	136	133
15	123	133	133
16	117	127	123
17	112	122	120
18	105	115	115
19	112	122	118
20	98	108	103
21	100	110	108

In order to infer whether the relative positions of individuals are significantly affected by the subjective probability technique, tables were set up to calculate Spearman's rank order coefficient for each of the sets of scores for both grades 11 and 12. The results are shown on Table VIII.

With 20 pairs of scores the Spearman coefficient must be larger than .377 to be significantly different from zero at the 5% level of confidence or more than .534 at the 1% level.

TABLE VIII

SPEARMAN'S CORRELATION RELATING I.Q.'S

Grade	Scores	r_s
11	Conventional - S.P. I.Q.	.807
11	S.P. I.Q. - Systematic	.974
12	Conventional - S.P. I.Q.	.867
12	S.P. I.Q. - Systematic	.967

In all cases the Spearman coefficient is above the required value for the 1% level of confidence which is .534. Like the DAT tests, the comparison of correlations by the Fisher transformation indicated that the differences were not significant at the 1% level of confidence.

For comparisons with other tests taken, by the same students, T-scores are provided for all I.Q. test results in Table IX.

TABLE IX

T-SCORES OF I.Q.'S

	Grade 11			Grade 12		
Student	I.Q.	S.P.I.Q.	Systematic I.Q.	I.Q.	S.P.I.Q.	Systematic I.Q.
1	62	70	70	71	68	51
2	70	63	63	58	68	68
3	48	63	63	54	63	68
4	65	60	56	49	61	57
5	59	55	60	54	59	57
6	55	55	56	62	58	59
7	48	54	52	59	56	63
8	52	51	56	62	54	53
9	56	49	50	66	54	52
10	50	49	49	51	53	59
11	38	48	48	56	52	51
12	44	42	40	52	51	54
13	40	38	38	44	49	48
14	59	58	52	41	49	50
15	55	52	52	53	48	47
16	51	46	46	50	47	48
17	45	44	45	41	46	46
18	42	40	44	47	44	44
19	45	44	42	47	44	44
20	30	35	30	38	42	41
21	35	30	35	34	41	42
22				38	39	37
23				44	34	34
24				47	37	39
25				29	29	29

A summary of the results of the teacher-prepared diagnostic tests with the calculation of Spearman's rank order correlation are included in Table X.

TABLE X

SPEARMAN'S CORRELATION RELATING RESULTS OF DIAGNOSTIC TESTS

Grade	Subject	r_s
10	Physics	.865
10	Geography	.779
11	Geography	.909
11	Chemistry	.906
12	Chemistry	.981
12	Physics	.997

In all cases the difference from zero is significant at 1% level.

The next stage in the study involved examining whether the subjective probability technique could contribute to more accurate predictions of future achievement. To do this it was decided to calculate the Pearson product moment between the results of each of the diagnostic tests and the mid-term examinations which are composed principally of essays and problems. Tables XXXIII through XXXVIII in the Appendix F provide the calculations of the Pearson product moments between the appropriate subject examinations and the two sets of scores on the diagnostic tests.

The difference between r_{sp} and r_{mc} was tested by Fisher's Z_r transformation. As this involved the unit-normal-curve, deviate values of 1.96 and 2.58 are required for significance at the 1% and 5% levels. As indicated by the table below, in no cases was the 1% reached.

TABLE XI

COMPARISON OF CORRELATIONS BETWEEN RESULTS OF DIAGNOSTIC TESTS AND RELATED MID-TERM EXAMINATIONS

Grade	Subject	r_{mc}	r_{sp}	Zr_{mc}	Zr_{sp}	Z
10	Physics	.411	.416	.437	.443	.005
10	Geography	.372	.462	.390	.500	.290
11	Geography	.686	.726	.840	.920	.138
11	Chemistry	.550	.556	.618	.628	.031
12	Chemistry	.695	.701	.858	.869	.031
12	Physics	.635	.884	.750	1.309	1.162

II. ANALYSIS OF CLASS KNOWLEDGE

The next stage of the study involved an item by item examination of a random selection of the tests to diagnose deficiencies in class knowledge and suggest a general review program. Ten names were selected at random from names in a course and Tables XII - XVII constructed to enable analysis of class knowledge. Included in them is a complete breakdown of the type of scoring produced by students who:

(a) had complete knowledge and confidence;

(b) have incomplete knowledge and corresponding confidence;

(c) have knowledge, but lacked confidence;

(d) have no knowledge and realize it;

(e) have no knowledge, but guess at the answer.

For purposes of this analysis, the following cause-effect relationships were assumed:

(a) 5/0 with 5 on the correct answer implies an understanding of the concept or knowledge of the information asked.

(b) distribution of weights among the choices with a major portion on the correct answer indicates partial confidence in their facts, which was taken to infer partial knowledge.

(c) 1/1/1/1/1 distribution implies no confidence in the ability to select the correct answer which was taken to infer a lesser degree of knowledge than in (b). This may show that they know nothing of these alternatives.

(d) total weight or a major portion placed on an incorrect alternative indicates misplaced confidence which was taken to infer misinformation or faulty concepts (i.e., partial knowledge of lesser degree than in (b)).

In each of the latter three cases the need for selective review or reteaching would be implied. If a major portion of the class responded in such a way, a reappraisal of teaching techniques may be considered; however, for individual cases, directed review may be in order.

TABLE XII

FREQUENCIES OF WEIGHTINGS OF ITEMS IN GRADE X PHYSICS

Score		Items 1	2	3	4	5	6	7	8	9	10	11
5/0	S_5	4	9	6	2	6	2	4	9		6	10
	S_0	4		3	6	1	5	4	1	8	4	
4/1	S_4					1						
	S_1											
	S_0	1			1			1				
3/2	S_3	1	1		1							
	S_2									1		
	S_0			1		1	1					
3/1/1	S_3											
	S_1						2					
	S_0											
2/2/1	S_2											
	S_1											
	S_0					1						
2/1/1/1	S_2											
	S_1											
	S_0											
1 =	S_1							1		1		

TABLE XII (Continued)

Score		Items 12	13	14	15	16	17	18	19	20	Totals
5/0	S_5	6	8	2	4	5	5	6	7	4	105
	S_0	2	2	5	4	2	1	1	2	4	59
4/1	S_4				1	1	2				5
	S_1	1									1
	S_0										3
3/2	S_3	1		1	1	1	2				9
	S_2							1		1	3
	S_0			1				1			5
3/1/1	S_3							1			1
	S_1										2
	S_0										
2/2/1	S_2										
	S_1										
	S_0										1
2/1/1/1	S_2										
	S_1										
	S_0										
1 =	S_1			1		1			1	1	6

TABLE XIII

FREQUENCIES OF WEIGHTING OF ITEMS IN GRADE X GEOGRAPHY

Score		Items 1	2	3	4	5	6	7	8	9	10	11
5/0	S_5	6		9	8	5	7	9	6	2	3	3
	S_0		2			1	1		1	6	4	3
4/1	S_4											
	S_1		1						1			
	S_0									1	1	
3/2	S_3	3				2					1	1
	S_2	1	2		1	1						
	S_0								1	1		1
3/1/1	S_3		1			1	1	1				
	S_1											
	S_0											
2/2/1	S_2						1					1
	S_1											
	S_0		1									
2/1/1/1	S_2											
	S_1										1	1
	S_0											
1/1/1/1/1	S_1		3	1	1			1				

TABLE XIII (Continued)

Score		Items 12	13	14	15	16	17	18	19	20	Totals
5/0	S_5	4		8	4	5	1	3	6	2	91
	S_0	1	5		3		2	4		2	35
4/1	S_4	1		1							2
	S_1										2
	S_0		1							1	4
3/2	S_3	2				2	1	1			13
	S_2	1			1		1				8
	S_0		1					1	1	3	9
3/1/1	S_3					1	1	1			7
	S_1								1		1
	S_0		1								1
2/2/1	S_2										2
	S_1										0
	S_0						1				2
2/1/1/1	S_2	1				1					2
	S_1			1					1	2	6
	S_0										
1/1/1/1/1	S_1		2		2	1	3		1		15

TABLE XIV

FREQUENCIES OF WEIGHTING OF ITEMS IN GRADE XI GEOGRAPHY

Score		Items 1	2	3	4	5	6	7	8	9	10	11
5/0	S_5	3	2	6	3	2	1		5			1
	S_0	2	2		1	1	2	3	3	1	3	4
4/1	S_4					1					1	1
	S_1						2					
	S_0		2		1							
3/2	S_3			2		1				1		
	S_2		1		1	1	1	1		2	1	2
	S_0	2	2		1	1				1	1	1
3/1/1	S_3	1					1	1				
	S_1			1			1	1		2		
	S_0											
2/2/1	S_2				1	1	1	1		1		
	S_1											
	S_0								1			
2/1/1/1	S_2										1	
	S_1							1				
	S_0											
1 -	S_1	1			1	1		1		1	2	

TABLE XIV (Continued)

Score		Items 12	13	14	15	16	17	18	19	20	Totals
5/0	S_5		2	3	2		6	1	3		40
	S_0	3	1	3	1	1	1	2	2	2	38
4/1	S_4							1	1		5
	S_1		1		1			1			5
	S_0									1	4
3/2	S_3		2	2	4	1	1	1	2		17
	S_2		3	1				2	1	1	18
	S_0	4				1	1	1			16
3/1/1	S_3									1	4
	S_1										5
	S_0					1					1
2/2/1	S_2										5
	S_1				1						1
	S_0	1									2
2/1/1/1	S_2										1
	S_1										1
	S_0										
1 =	S_1	1				5				4	17

TABLE XV

FREQUENCIES OF WEIGHTING OF ITEMS IN GRADE XI CHEMISTRY

Score		Items 1	2	3	4	5	6	7	8	9	10	11
5/0	S_5	9	9	9	6	5	10	6	5	9	7	5
	S_0					2		3	3	1		2
4/1	S_4	1		1	1	2			1		2	2
	S_1								1			1
	S_0		1									
3/2	S_3				1	1		1			1	
	S_2				1							
	S_0				1							
3/1/1	S_3											
	S_1											
	S_0											
2/2/1	S_2											
	S_1											
	S_0											
2/1/1/1	S_2											
	S_1											
	S_0											
1 =	S_1											

TABLE XV (Continued)

Score		Items 12	13	14	15	16	17	18	19	20	Totals
5/0	S_5	7	9	8	9	6	10	8	6	10	153
	S_0	2	1	2		2		2	2		22
4/1	S_4										10
	S_1								1		3
	S_0										1
3/2	S_3				1				1		6
	S_2					1					2
	S_0	1				1					3
3/1/1	S_3										
	S_1										
	S_0										
2/2/1	S_2										
	S_1										
	S_0										
2/1/1/1	S_2										
	S_1										
	S_0										
1 =	S_1										

TABLE XVI

FREQUENCIES OF WEIGHTING OF ITEMS IN GRADE XII CHEMISTRY

Score		Items 1	2	3	4	5	6	7	8	9	10	11
5/0	S_5	9	7	5	7	4	5	6	4	5	6	4
	S_0	1	1			4	4	1	4		1	2
4/1	S_4				2	1		1		3		3
	S_1			1								
	S_0		1	2	1			1	2		1	
3/2	S_3		1	1				1		1		
	S_2						1				2	
	S_0			1								1
3/1/1	S_3											
	S_1									1		
	S_0											
2/2/1	S_2											
	S_1											
	S_0											
2/1/1/1	S_2											
	S_1											
	S_0											
1 =	S_1					1						

TABLE XVI (Continued)

Score		Items 12	13	14	15	16	17	18	19	20	Totals
5/0	S_5	8	1	1	4	3	7				86
	S_0		1	6	2		3				30
4/1	S_4		1	2	1	1					15
	S_1		2								3
	S_0	1	1			1					11
3/2	S_3	1									5
	S_2		1			4					8
	S_0										2
3/1/1	S_3		1		2						3
	S_1										1
	S_0				1						1
2/2/1	S_2		1								1
	S_1		1								1
	S_0			1							1
2/1/1/1	S_2										
	S_1										
	S_0										
1 =	S_1					1					2

TABLE XVII

FREQUENCIES OF WEIGHTING OF ITEMS IN GRADE XII PHYSICS

Score		Items 1	2	3	4	5	6	7	8	9	10	11
5/0	S_5	8	7	6	6	5	7	6	4	4	5	9
	S_0	2	2	2		2	1	2	3	2	2	
4/1	S_4								1			
	S_1											
	S_0											
3/2	S_3		1	1	1					1	1	
	S_2								1			
	S_0				1	1				1		
3/1/1	S_3									1		
	S_1											
	S_0											
2/2/1	S_2											
	S_1											
	S_0											
2/1/1/1	S_2				1							
	S_1										1	
	S_0											
1 =	S_1			1	1	2	2	2	1	1	1	1

TABLE XVII (Continued)

Score		12	13	14	15	16	17	18	19	20	Totals
		Items									
5/0	S_5	6	7	4	7	2	6	2	4	7	112
	S_0	1	1	2	3	7	1	8	6	3	50
4/1	S_4		1								2
	S_1										0
	S_0						1				1
3/2	S_3	1		1							7
	S_2			1							2
	S_0					1					4
3/1/1	S_3										1
	S_1			1							1
	S_0										
2/2/1	S_2										
	S_1										
	S_0										
2/1/1/1	S_2										1
	S_1										1
	S_0										
1 =	S_1	2	1	1			2				18

By examining the tables of the weighting of items on the various tests it is possible to arrive at suggestions for application of the diagnostic information.

Many of the items on the Grade X Physics Test have been scored in the 5/0 method with 59 of the 164 indicating a misconception. Reteaching of a concept is indicated from the results of items 4, 6, 9 and 14. It may be in order to question the class to verify that they were not guessing rather than using the 1/1/1/1/1 selection.

By the scattering of scoring methods with inappropriate weighting on the Grade X Geography Test, a lack of confidence in the material tested by items 2, 11, 12, 17 and 20 is indicated. A review of these topics should be provided. The reteaching of the material tested by item 9 is indicated.

The inappropriate placing of weights on the Grade XI Geography Test on items 2, 6, 7, 9, 11, 12 and 18 indicates the need for review of these areas. The material tested by items 16 and 20 requires reteaching.

Although items 7 and 8 represent weak areas in the Grade XI Chemistry Test, the material covered seems to be adequately learned and understood. No major review work is indicated.

More than half of the students selected an inappropriate choice for items 6 and 8 on the Grade XII Chemistry Test. This would indicate the necessity of reteaching the material tested by these items.

The inappropriate weighting on items 16, 18 and 19 of the Grade XII Physics Test indicate the necessity of reteaching the material tested.

III. ANALYSIS OF INDIVIDUAL'S DEFICIENCIES

The final stage of the study involved selecting a single student's results for each test and examining them item by item to diagnose weaknesses and to suggest specific steps to take in improving his future performance. Tables XVIII through XXIII provide this information for six, randomly selected students.

The student selected from the Grade X Physics class shows a lack of confidence in eight questions and only partial confidence in three more, indicating the desirability of reinforcement by way of reviewing each of these areas. In only three questions was he misinformed--i.e., sure of a wrong answer and thus reteaching as such would be at a minimum.

The student in Grade X Geography has complete information on only five of the twenty items. A minor review is indicated by items 1, 4, 6, 11 and 20, while complete reteaching is indicated for items 2, 8, 9, 10, 13, 16 and 17. Items 15 and 18 indicate misinformation and will require special attention.

Review of items 2, 6, 7, 9, 10, 11, 12 and 20 would build up the confidence and knowledge for the student of Grade XI Geography. Item 16 requires reteaching.

The Grade XI Chemistry student requires very little reteaching. Items 12 and 18 were missed entirely and indicate a lack of understanding of the concept tested while review of items 4 and 16 is indicated.

Reteaching of items 2, 5 and 17 is indicated for the student of Grade XII Chemistry. A review of items 3, 6, 7, 13, 15 and 16 is required to complete knowledge in these areas while review of items 9 and 11 should be instituted to reinforce the confidence.

The Grade XII Physics student has misinformation of items 9, 15, 16, 18, 19 and 20 thus requiring reteaching of these area. The lack of confidence in items 10, 12, 13 and 14 indicates the desirability of review work here.

It should be emphasized at this point that these represent a sampling of students' results and are meant merely as a pattern to indicate the means of assessing individual scores. In practice one would either assess all results if time were available, or if not, he would select those which collaborating evidence indicate require further examination or a more complete diagnoses--i.e., the results of the subjective probability scoring could be considered of a corroborating nature.

TABLE XVIII

AN INDIVIDUAL'S WEIGHTING OF ITEMS IN GRADE X PHYSICS

Item	Weighting	S.P.	Conventional
1	3/2 S_3	.84	1
2	3/2 S_3	.84	1
3	3/2 S_0	.74	0
4	5/0 S_0	0	0
5	3/2 S_0	.24	0
6	3/2 S_0	.24	0
7	5/0 S_0	0	0
8	5/0 S_5	1	1
9	3/2 S_2	.64	0
10	5/0 S_5	1	1
11	5/0 S_5	1	1
12	3/2 S_3	.84	1
13	5/0	1	1
14	3/2	.84	1
15	3/2	.84	1
16	3/2	.84	1
17	3/2	.84	1
18	3/2 S_0	.24	0
19	5/0 S_5	1	1
20	5/0	0	0
		11.4	12 - 2 = 10

TABLE XIX

AN INDIVIDUAL'S WEIGHTING OF ITEMS IN GRADE X GEOGRAPHY

Item	Weighting	S.P.	Conventional
1	3/2 S_2	.64	0
2	1 = S_1	.60	0
3	5/0 S_5	1.0	1
4	3/2 S_2	.64	0
5	5/0 S_5	1	1
6	2/2/1 S_2	.72	0
7	5/0 S_5	1	1
8	4/1 S_1	.36	0
9	4/1 S_0	.16	0
10	4/1 S_0	.16	0
11	2/2/1 S_2	.72	1
12	5/0 S_5	1	1
13	1 = S_1	.6	0
14	4/1 S_4	.96	1
15	5/0 S_0	0	0
16	1 = S_1	.6	1
17	1 = S_1	.6	0
18	S/0 S_0	0	0
19	S/0 S_5	1	1
20	2/111 S_1	.56	0
		12.3	8

TABLE XX

AN INDIVIDUAL'S WEIGHTING OF ITEMS IN GRADE XI GEOGRAPHY

Item	Weighting	S.P.	Conventional
1	5/0	1	1
2	3/2 S_2	.64	0
3	5/0	1	1
4	5/0	1	1
5	5/0	1	1
6	3/2 S_2	.64	0
7	2/2/1 S_2	.52	1
8	5/0	1	1
9	3/2 S_3	.84	1
10	2/111 S_2	.72	1
11	4/1 S_4	.96	1
12	3/2 S_0	.24	0
13	5/0 S_5	1	1
14	5/0	1	1
15	5/0	1	1
16	1 =	.6	0
17	5/0	1	1
18	3/2 S_3	.84	1
19	5/0	1	1
20	3/2 S_2	.64	0
		16.6	15 - 1 = 14

TABLE XXI

AN INDIVIDUAL'S WEIGHTING OF ITEMS IN GRADE XI CHEMISTRY

Item	Weighting	S.P.	Conventional
1	5/0 S_5	1	1
2	5/0 S_5	1	1
3	5/0 S_5	1	1
4	3/2 S_2	.64	0
5	5/0 S_5	1	1
6	5/0 S_5	1	1
7	5/0 S_5	1	1
8	5/0 S_5	1	1
9	5/0 S_5	1	1
10	5/0 S_5	1	1
11	5/0 S_5	1	1
12	5/0 S_0	0	0
13	5/0 S_5	1	1
14	5/0 S_5	1	1
15	5/0 S_5	1	1
16	3/2 S_0	.24	0
17	5/0 S_5	1	1
18	5/0 S_0	0	0
19	5/0 S_5	1	1
20	5/0 S_5	1	1

TABLE XXII

AN INDIVIDUAL'S WEIGHTING OF ITEMS IN GRADE XII CHEMISTRY

Item	Weighting	S.P.	Conventional
1	5/0 S_5	1	1
2	5/0 S_0	0	0
3	4/1 S_0	.16	0
4	5/0 S_5	1	1
5	S_0	0	0
6	3/2 S_2	.64	0
7	4/1 S_0	.16	0
8	5/0 S_5	1	1
9	4/1 S_4	.96	1
10	3/2 S_2	.64	0
11	4/1 S_4	.96	1
12	5/0 S_5	1	1
13	3/2 S_2	.64	0
14	5/0 S_5	1	1
15	3/1 S_1	.28	0
16	S_0	.16	0
17	5/0 S_0	0	0

TABLE XXIII

AN INDIVIDUAL'S WEIGHTING OF ITEMS IN GRADE XII PHYSICS

Item	Weighting	S.P.	Conventional
1	5/0 S_5	1	1
2	5/0 S_5	1	1
3	5/0 S_5	1	1
4	5/0 S_5	1	1
5	5/0 S_5	1	1
6	5/0 S_5	1	1
7	5/0 S_5	1	1
8	5/0 S_5	1	1
9	5/0 S_0	0	0
10	3/2 S_3	.84	1
11	5/0 S_5	1	1
12	1 =	.6	1
13	1 =	.6	1
14	1 =	.6	1
15	5/0 S_5	1	1
16	5/0 S_5	0	0
17	5/0 S_5	1	1
18	5/0 S_0	0	0
19	5/0 S_0	0	0
20	5/0 S_0	0	0

IV. STUDENTS' REACTIONS

The results of the survey of grade 10 student reaction to the new technique of testing is as follows:

Question 1 - Which form of test do you prefer?

Answer (a) -	Conventional multiple-choice	6
(b) -	Subjective probability	30

Question 2 - If you could arbitrate the type of scoring to be used on your mid-term exam, which of the following would you select?

Answer (a) -	All subjective probability items	21
(b) -	Some subjective probability items	9
(c) -	No items of subjective probability type	6

V. SUMMARY

The study is a comparison of conventional and subjective probability results on standardized tests. National norms were used to establish I.Q. scales and percentile ranks. A comparison of achievement within a given group is provided by T-scores.

The Spearman correlations were calculated to indicate what effect on individual rank in a class the subjective probability technique has. The Pearson coefficient provides a means of indicating relative value of the subjective probability technique in predicting success on essay and problem type examinations.

The analysis of classes' results provide a means of planning remedial action with a class because of more insight into the reason for errors than would be the case with the conventional test results. As required, individual results can be analysed to provide specific information for individual remedial action in any given subject.

CHAPTER V

CONCLUSION AND EVALUATION

The findings of this study indicate that there is no significant difference in the rank of a student in class when using the subjective probability method rather than the conventional one of scoring a test. This observation was found true both for the standardized tests of the Differential Aptitude Battery and the Otis Mental Ability as well as for the tests which had been prepared specifically for purposes of diagnosing weakness by means of the subjective probability technique. While there was a slightly high correlation between subjective probability scores and those of essay type examinations than there was between the conventional multiple-choice score and the essay type examination, the Fisher transformation did not exceed 1.96.

The first "null" hypothesis was rejected, and the alternative hypothesis accepted:

1. "The Spearman rank order correlation between the results of conventional and subjective probability treatment does differ significantly from zero."

The second null hypothesis was accepted as stated:

2. "The Pearson product moment correlation between conventionally treated subject tests and mid-term examination results does not differ significantly from the

Pearson product moment correlation between subjective probability treated subject tests and mid-term examination results."

TABLE XXIV

SPEARMAN'S RANK ORDER CORRELATION RELATING SCORING TECHNIQUES

Test	r_s
DAT - VR	.589## -- .669#
DAT - AR	.554## -- .830#
Otis	.807# -- .974#
10 Physics	.865 #
10 Geography	.779 #
11 Geography	.909 #
11 Chemistry	.906 #
12 Chemistry	.981 #
12 Physics	.997 #

1% level of confidence. ## 5% level of confidence.

TABLE XXV

FISHER'S TRANSFORMATION COMPARING PEARSON PRODUCT MOMENTS

Grade	Subject	Z
10	Physics	.005
10	Geography	.290
11	Geography	.138
11	Chemistry	.031
12	Chemistry	.031
12	Physics	1.162

As these are based on the normal curve, values of at least 1.96 would be required to indicate a difference at the 1% level of confidence.

It was recognized that if the first hypothesis were to be rejected, the second would necessarily be accepted. If, however, there had been no significant correlation between the ranks under the two scoring techniques, it would have been possible for the diagnostic test scores provided by one technique to be more useful for predictive purposes than those provided by the other technique.

The consequence of the accepted hypotheses is that no information about a student's knowledge is lost by using the subjective probability technique rather than the conventional one, and much valuable additional information is available. This additional information, in the form of indicated degrees of partial confidence, provides an efficient tool for diagnosing deficiencies in a student's knowledge or skill and makes it possible to channel efforts toward the areas in which remedial work is required.

In this study, all items were of the type which had only one correct answer. In the standardized tests the other alternatives serve merely as distractors. In the diagnostic items each alternative was designed in order that the type of misconception involved in its selection could be identified. A possible extension of the subjective probability procedure would be to develop items with gradations of acceptability--i.e., a>b>c>d>e. In such an item the alternatives would be ranked rather than weighted, and a different type of scoring formula would have to be developed.

Although the standardized tests could not be used to suggest any procedure for reteaching, they did supply the same separation of students when used in the subjective probability as in the conventional manner. These were tests of known reliability which appeared to apply both to the conventional and to the subjective probability approach.

A possible extension of the application of this testing technique would be to set up a complete program in a given subject area and have test items prepared for every stage of progress. These items could be handled by a type of mechanical or electronic recording machine which would automatically record the individual weighting or ranking and supply both a profile of class results and individual choices. Such a program would supply continuous individual remedial teaching without introducing a programmed learning process.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Adams, Georgia Sacks in consultation with T.L. Torgerson. Measurement and Evaluation in Education, Psychology and Guidance. Atlanta: Holt Rinehart and Winston, 1954.

2. Bennett, George K.; Seashore, Harold G.; and Wesman, Alexander G. Differential Aptitude Tests. New York: The Psychological Corporation, 1962.

3. Bennett, George K.; Seashore, Harold G.; and Wesman, Alexander G. Differential Aptitude Tests Manual. 4th ed. New York: The Psychological Corporation, 1966.

4. Cohen, John. "Subjective Probability," Scientific American, CXCVII (Nov., 1957), 128-138.

5. Cohen, John; Dearnaley, E.J.; and Hansel, C.E.M. "Measures of Subjective Probability," British Journal of Psychology, XLVIII (1957), 271-275.

6. Coombs, C.H. "On the Use of Objective Examinations," Educational and Psychological Measurement, XIII (1953), 308-310.

7. Coombs, C.H.; Mulholland, T.; and Womer, F.B. "Assessment of Partial Knowledge," Educational and Psychological Measurement, XVI (1956), 13-37.

8. de Finetti, Bruno. "Methods for Discriminating Levels of Partial Knowledge Concerning a Test Item," British Journal of Mathematics and Statistical Psychology, XVIII (1965), 87-123.

9. Ferguson, George A. Statistical Analysis in Psychology and Education. New York: McGraw Hill, 1959.

10. Flavell, John H. The Developmental Psychology of Jean Piaget. Princeton: Van Nostrand, 1963.

11. Furst, Edward J. Constructing Evaluation Instruments. New York: Longmans, Green and Company, 1958.

12. Graesser, R.F. "A Note on Multiple Choice Tests," School Science and Mathematics Journal, LXVI (Nov., 1966), 526.

13. Gritten, Frances, and Johnson, D.M. "Individual Differences in Judging Multiple Choice Questions," Journal of Educational Psychology, XXXII (1941), 423-30.

14. Hansel, C.E.M. "The Addition of Subjective Probabilities," Acta Psychologica, XV (1956), 371-80.

15. Jenkins, Susan E. "The Use of the Subjective Probability Test Method with Fourth Grade Students." Unpublished M.Ed. thesis, Rutgers University, 1967.

16. Kyburg, H.E. Jr., and Smokler, H.E., eds. Studies in Subjective Probability. New York: Wiley, 1962.

17. Leake, L. Jr. "The Status of Three Concepts of Probability in Children of Seventh, Eighth and Ninth Grades," Journal of Experimental Education, XXXIII (1965), 34-72.

18. Otis, Arthur S. Otis Quick Scoring Mental Ability Tests--Gamma Test: Form AM. New York: Harcourt, Brace and World Inc., 1937.

19. Shuford, Emir H., and Massengill, H.E. The Effects of Guessing on the Quality of Personnel and Counselling Decisions: The First Semiannual Technical Report Under Contract AF49(638)-1744. ARPA Order number 833, Lexington: Shuford-Massengill Corp., 1966.

20. Taylor, Peter A., and Jenkins, S.E. "Subjective Probability Testing with Fourth Grade Students," Manitoba Journal of Educational Research, IV (June, 1969), 20-32.

21. Van der Meer, H.C. "Decision Making: Need for Achievement and Probability Preference Under Chance and Skill Orientation," Acta Psychologica, XXVI (1967), 353-72.

22. Van Naersen, R.F. "A Scale for the Measurement of Subjective Probability," Acta Psychologica, XX (1961), 159-166.

23. Yost, Patricia A.; Siegel, A.E.; and Andrews, J. "Non-Verbal Probability Judgment by Young Children," Journal of Child Development, XXXIII (Dec., 1962), 767-80.

APPENDIX

APPENDIX A

DIAGNOSTIC TEST PAPERS

GRADE 10 PHYSICS

1. A man pushes a puck on a frictionless horizontal surface with a force of 10 newtons. The resulting acceleration is 4.0 meters/second^2. What is the mass of the puck?

   A. 0.4 kg
   B. 2.5 kg
   C. 4.0 kg
   D. 10 kg
   E. 40 kg

2. In ALL EXCEPT ONE of the following situations, an object is being accelerated. Which one is the exception?

   A. The object changes direction without changing speed.
   B. The object changes speed without changing direction.
   C. The object maintains speed and direction.
   D. The object maintains uniform circular motion.
   E. The object moves in the trajectory of a projectile.

Questions 3 and 4 refer to the following situation:
During a planned maneuver in sapce flight, a free-floating astronaut pushes a free-floating instrument package. The mass of the astronaut is greater than that of the instrument package.

3. The force exerted by the astronaut on the package

   A. is equal to the force exerted by the package on the astronaut.

B. is greater than the force exerted by the package on the astronaut.

C. is less than the force exerted by the package on the astronaut.

D. is equal to zero.

E. may be greater than, less than, or equal to the force exerted by the package on the astronaut; one cannot tell with the information given here.

4. During the push

A. the magnitude of the acceleration of the astronaut is greater than that of the instrument package.

B. the magnitude of the acceleration of the astronaut is smaller than that of the instrument package.

C. neither astronaut nor instrument package are accelerated.

D. the accelerations of each are equal in magnitude but opposite in direction.

E. the accelerations of each are equal in magnitude and equal in direction.

Questions 5 and 6 refer to the following statement.

Scientists on the imaginary planet Q have defined a unit of length, the "lar", to be the distance between two mountain peaks on the surface of the planet. The unit of time on the planet Q is called the "tik", and is defined as the average interval between pulse beats of the king.

5. What units would express acceleration on planet Q if acceleration were defined as it is on earth?

A. lar/tik
B. lar/sec
C. lar^2/sec
D. tik/lar^2
E. lar/tik^2

6. If the distance between the cities Zytropolis and Elany on planet Q is 20 lars, what would your average speed be if you made the trip in 100 tiks?

A. 0.2 lars/tik
B. 0.1 tiks/lar
C. 5 tiks/lar
D. 5 lars/tik
E. 100 tiks/lar

Questions 7 and 8 refer to the following statements concerning Galileo's work with balls rolling down inclined planes. This work led to the acceptance of the idea that falling objects accelerate uniformly. The quotations are from Two New Sciences.

1. If speed during fall increases with time, $\frac{d}{t^2}$ is constant.
2. "We took a piece of wooden scantling, about 12 cubits long, half a cubit wide, and three finger breadths thick. In its top edge we cut a straight channel."
3. "Having raised the scantling in a sloping position by raising one end some one or two cubits above the other, we let the ball roll down the channel."
4. "We always found that the distances traversed were to each other as the squares of the times."

5. Since for a rolling ball $\frac{d}{t^2}$ is constant, $\frac{d}{t^2}$ is constant for a falling ball also.

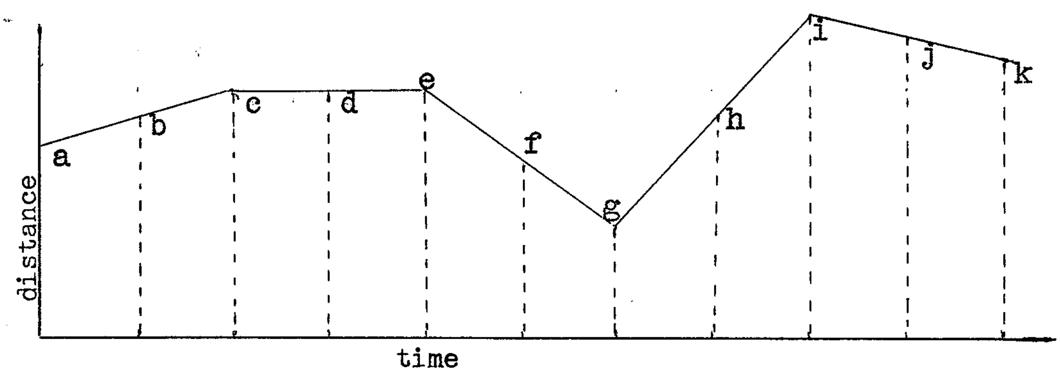
7. Which of the statements are assumptions made by Galileo?

A. 1 only
B. 4 only
C. 5 only
D. 1 and 4 only
E. 4 and 5 only

8. Which statement presents experimental results?

A. 1 only
B. 4 only
C. 5 only
D. 2 and 3 only
E. 1 and 5 only

Questions 9 and 10 refer to the following graph.



9. The greatest distance is traveled between the times corresponding to points

A. a and c
B. c and e
C. e and g
D. g and i
E. i and k

10. v is greatest between times corresponding to points

A. a and k

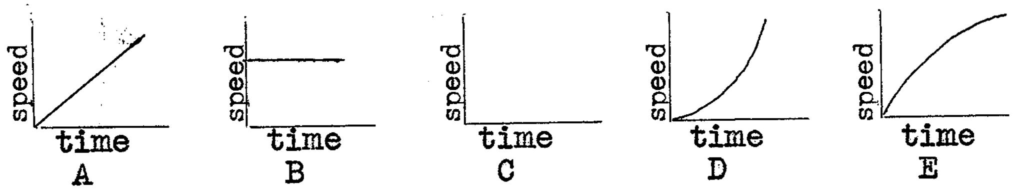
B. c and e

C. c and g

D. g and i

E. i and k

11. A cart, initially at rest, is pulled with a constant, unbalanced force. Which graph best represents how the speed of the cart changes with time?



A. graph A

B. graph B

C. graph C

D. graph D

E. graph E

Questions 12 and 13 refer to the following statement and table.

Main Street in Centerville is crossed by streets called 1st Street, 2nd Street, 3rd Street ..... 46th Street. Blocks between the numbered streets are equally long.

Five cars are travelling along Main Street, and their locations are recorded at five-minute intervals, as shown in the following table:

	10:00 a.m.	10:05 a.m.	10:10 a.m.	10:15 a.m.
car A	15th	20th	25th	30th
car B	30th	25th	15th	10th
car C	1st	2nd	5th	10th
car D	9th	10th	20th	38th
car E	35th	33rd	23rd	20th

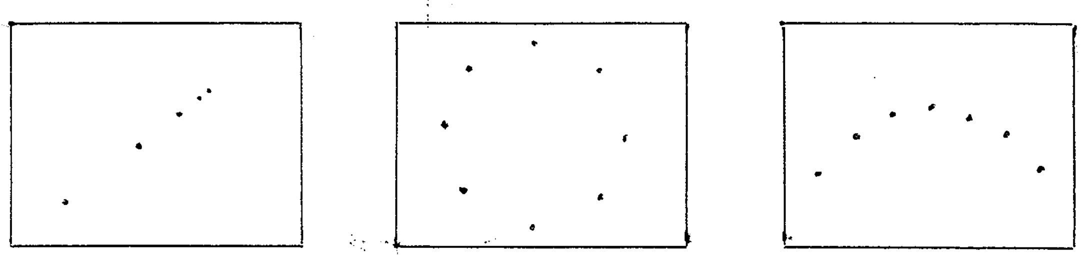
12. Which car traveled with the greatest average speed during the period described?

A. car A
B. car B
C. car C
D. car D
E. car E

13. Assuming that all cars started from est at 10:00 a.m., which car could have traveled with uniform acceleration during the entire period described?

A. car A
B. car B
C. car C
D. car D
E. car E

For Questions 14 and 15 use the following figures which represent stroboscopic photographs of a moving ball. The strobe rate is constant and is the same for all three "photographs".



14. Which of the "photographs" could have been produced with the camera in motion and the ball fixed in position?

A. none
B. 1 only
C. 2 only
D. 1 and 2 only
E. 1, 2 and 3

15. If the camera is fixed in position, which of the "pictures" show a ball being acted upon by a net unbalanced force?

A. 1 only
B. 3 only
C. 1 and 3 only
D. 2 and 3 only
E. 1, 2 and 3

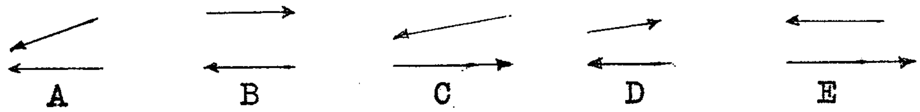
16. Which of the following increases with time if an object moves with uniform velocity?

A. instantaneous velocity
B. average velocity
C. acceleration
D. direction
E. displacement

17. A sprinter reaches top speed 3 seconds after the start of a race. In those 3 seconds, he moves 18 meters. Assume that he accelerates uniformly. What is his acceleration?

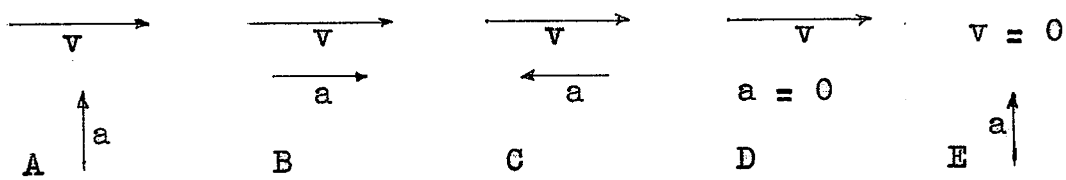
A. 2 meters/sec^2
B. 3 meters/sec^2
C. 4 meters/sec^2
D. 7 meters/sec^2
E. 18 meters/sec^2

18. Which of the following pairs of vectors has the resultant of largest magnitude?



A. pair A
B. pair B
C. pair C
D. pair D
E. pair E

19. The arrows in the diagrams below show the directions of the velocity and acceleration vectors that apply to a car at five separate instants of time.



ALL EXCEPT ONE of the diagrams above show an instant at which the velocity is changing. Which is the exception?

A. diagram A
B. diagram B
C. diagram C
D. diagram D
E. diagram E

20. A satellite is in a circular orbit around the earth. Which of the following statements must be true?

1. The speed is constant.
2. The velocity is constant.
3. The period is constant.

A. 1 only
B. 2 only
C. 1 and 3 only
D. 2 and 3 only
E. 1, 2 and 3

GRADE 10 GEOGRAPHY

1. Indicate the term which is most closely related to the underlined term: scale:

   A. distance
   B. ratio
   C. time
   D. angle
   E. pressure

2. Large scale maps show:

   A. great area in great detail
   B. small area in great detail
   C. great area in little detail
   D. small area in little detail
   E. great area with no detail

3. The number of degrees of longitude in one hour is:

   A. 50
   B. 10
   C. 15
   D. 20
   E. 25

4. Till, boulder clay and eskers are associated with:

   A. conservation
   B. glaciation
   C. natural vegetation
   D. mining
   E. reforestation

5. The Maritimes are most famous for:

   A. cod fishing
   B. ship building
   C. iron and steel manufacturing
   D. lobster trapping
   E. agriculture

6. The Monteregian hills are:

A. buttes
B. escarpments
C. erosional remnants
D. metamorphic rocks
E. pene plains

7. Winnipeg's climate is classified as:

A. humid continental
B. subtropical
C. tundra
D. sub-arctic
E. humid marine

8. The St. Lawrence Great Lakes Lowlands is divided by:

A. The Niagara Escarpment
B. The Frontenac Axis
C. The Monteregian Hills
D. The Champlain Fault
E. The Great Lakes

9. Which of the following does not deal with conservation:

A. shifting cultivation
B. terracing
C. crop rotation
D. reforestation
E. wind breaks

10. Much of the Great Lakes St. Lawrence Lowlands are covered by:

A. loess
B. alluvial deposits
C. glacial drift
D. sand dunes
E. lagoons

11. Continentality in North America causes great extremes in temperatures. Which of the following areas has the most uniform temperature?

A. Winnipeg
B. Victoria
C. Ottawa
D. St. Johns
E. Niagara Peninsula

12. Chernozem soils are the most fertile on the North American continent. What kind of vegetation corresponds most closely to this soil type?

A. Boreal
B. Tundra
C. Arctic
D. Semi-desert
E. Grassland

13. The largest glacial delta is on which river?

A. Assiniboine
B. Ottawa
C. South Saskatchewan
D. Fraser
E. Red

14. Perhaps the future of Canadian development will be found in the important development of which of the following regions?

A. Prairies
B. St. Lawrence Lowlands
C. Appalachia
D. The North
E. West Coast

15. Palliser's triangle refers to

A. Niagara Peninsula
B. an area in Rocky Mountains
C. Lake of the Woods
D. the Peace River area
E. the dry prairie

16. The geologic structure of the prairie region can be best described as:

A. crystalline
B. sedimentary
C. folded
D. faulted
E. metamorphic

17. Mixed farming on the prairies is associated with

A. podzolic soils
B. brown soils
C. black soils
D. grey wooded soils
E. dark brown soils

18. The second most important grain crop on the prairies is

A. oats
B. wheat
C. rye
D. barley
E. flax

19. The term erosional remnant is used to describe the

A. Champlain Fault
B. Frontenac Axis
C. Laurentide Scarp
D. Monteregion Hills
E. Niagara Escarpment

20. The source region of hot air masses affecting Winnipeg in July and August is:

A. North Pacific
B. Pacific
C. Caribbean
D. Arizona
E. Montana

GRADE 11 GEOGRAPHY

1. The earth has neither of its poles tilted towards the sun at

   A. aphelion
   B. apogee
   C. equinox
   D. perigee
   E. perihelion

2. Which of the following features is associated with conic projections?

   A. equal area
   B. hachuring
   C. space grids
   D. standard parallel
   E. meridians and parallels to each other

3. Which of the following features of vulcanicity are instrusive in nature?

   A. calderas
   B. batholiths
   C. solfaturas
   D. mofettes
   E. geysers

4. Isostasy is related to

   A. theory of geosynclines
   B. continental drift
   C. vertical earth movements
   D. fold mountains
   E. horizontal earth movements

5. Which of the following terms is associated with mechanical weathering?

A. exfoliation
B. laterite
C. corrasion
D. corrosion
E. attrition

6. A river whose original course is determined by the lie of the land is said to be

A. subsequent
B. obsequent
C. consequent
D. braided
E. dendritic

7. Valley glaciers are responsible for the formation of certain physical features. Which of the following are not depositional in form?

A. aretes
B. drumlins
C. erratics
D. eskers
E. kames

8. Solution action by ground water is not responsible for which of the following features

A. stalactites
B. dolines
C. gorges
D. stalagmites
E. dry valleys

9. Which of the following coastal scenery features are not formed by submergence?

A. geos
B. drowned valleys
C. fiords
D. rias
E. elongated offshore islands

10. Wind is the most important erosional agent in arid regions. Which of the following features is found only in sandy deserts?

A. yardangs
B. inselbergs
C. depressions
D. seifs
E. overhangs

11. The position of the earth on its orbit when it is furthest away from the sun is called

A. apogee
B. aphelion
C. equinox
D. perigee
E. perihelion

12. Which of the following terms is not a feature associated with folding?

A. overthrusts
B. nappes
C. recumbent
D. isoclinal
E. graben

13. Of the following features which is not associated with vulcanism

A. sills
B. plugs
C. laccoliths
D. barchans
E. calderas

14. Which of the following materials is carbonaceous in nature?

A. coral
B. limestone
C. granite
D. anthracite
E. marble

15. When water is acting as a solvent or chemical in weathering of rocks, it is said to be wearing the rocks away by

A. corrosion
B. hydraulic action
C. corrasion
D. attrition
E. slumping

16. Deflation by the wind in arid regions is stopped by

A. hammada
B. erg
C. underground water
D. reg
E. yardangs

17. One of the following terms does not describe a feature of a river in its torrent tract.

A. rapids
B. waterfalls
C. meanders
D. potholes
E. rapid downcutting

18. Which of the following features is not caused by a lowering of sea level?

A. rejuvenation
B. rias
C. raised beaches
D. nick point
E. terraces

19. The word doline describes a feature of

A. chalk scenery
B. granite landscape
C. coastal scenery
D. limestone scenery
E. lacustrian landscape

20. Kettle holes are examples of

A. coastal erosion
B. coastal deposition
C. work of rivers
D. glacial erosion in valleys
E. glacial effects on lowlands

GRADE 11 CHEMISTRY

Test on Significant Figures

In Questions 1 - 20 below, state the number of significant figures in the number given.

1. 30916 — A. 1 B. 3 C. 4 D. 5 E. none of these

2. 70200 — A. 1 B. 2 C. 3 D. 4 E. 5

3. 80409 — A. 1 B. 3 C. 4 D. 5 E. none of these

4. 0.0007415 — A. 4 B. 7 C. 8 D. 9 E. none of these

5. 98120.0 — A. 3 B. 4 C. 5 D. 6 E. 7

6. 372.01020 — A. 2 B. 3 C. 5 D. 9 E. none of these

7. 734000. — A. 3 B. 5 C. 6 D. 7 E. none of these

8. 681.200 — A. 4 B. 5 C. 6 D. 7 E. none of these

9. 47201020 A. 3 B. 5
C. 6 D. 8 E. none of these

10. 20042 A. 1 B. 3
C. 4 D. 5 E. none of these

11. 36240. A. 3 B. 4
C. 5 D. 6 E. none of these

12. 0.0814283 A. 5 B. 6
C. 7 D. 8 E. none of these

13. 2010310 A. 1 B. 4
C. 5 D. 7 E. none of these

14. 1420.084 A. 3 B. 5
C. 6 D. 7 E. none of these

15. 28000 A. 1 B. 2
C. 3 D. 4 E. 5

16. 0.001020 A. 2 B. 4
C. 6 D. 7 E. none of these

17. 849.72 A. 3 B. 4
C. 5 D. 6 E. none of these

18. 0.32147 A. 4 B. 5
C. 6 D. 7 E. none of these

19. 600201.7 A. 1 B. 4
C. 6 D. 7 E. none of these

20. 3210800 A. 3 B. 4
C. 5 D. 6 E. 7

GRADE 12 CHEMISTRY

Multiple Choice Questions:

1. Samples of the following gases are measured under the same conditions of temperature and pressure. The gas with the highest density is:

   A. SO_2
   B. CH_4
   C. NH_3
   D. CO_2
   E. H_2S

2. Gases do not obey the ideal gas laws in exact detail. One of the reasons for this is that the gas molecules do have significant volumes. Deviations explanable by this reason are most apparent at:

   A. low pressures
   B. high temperatures
   C. high pressures
   D. low density
   E. low temperature and pressure

3. At 20° C, the vapour pressure of liquid chlorobenzene is 9 mm Hg pressure; at the same temperature the vapour pressure of alcohol is 44 mm Hg pressure. From this information we can deduce that:

   A. the critical temperature of chlorobenzene is lower than that of alcohol.

...... Continued

3\. B. the attractive forces between the chlorobenzene molecules are greater than those between the alcohol molecules.

C. at 600 mm Hg pressure chlorobenzene boils at a lower temperature than alcohol.

D. it is impossible for these two liquids to boil at the same temperature.

E. the vapour pressure of chlorobenzene at its normal boiling point is less than that for alcohol.

4\. Which of the following statements CANNOT be predicted from a knowledge of the kinetic theory of gases ONLY?

A. gases are easily compressible

B. gases are miscible

C. gases exert pressure

D. the molar volume of all gases at S.T.P. is 22.41

E. the pressure exerted by a gas increases as its temperature is increased.

5\. According to the kinetic theory, for a mixture of two ideal gases A and B:

A. All molecules have the same speed.

B. All molecules have the same kinetic energy.

C. The average molecular kinetic energy is greater for the heavier gas.

D. The average molecular speed is greater for the lighter gas.

.... Continued

5. E. The frequency of collision of A molecules with the container walls is affected by removing or adding more B molecules.

6. 2.00_g of N_2 are placed in a 5.00 litre container at 25^o C. The pressure of the gas is:

A. 0.350 atom

B. 1.75 atom

C. 0.700 atom

D. 5.25 atom

E. none of these

7. The behaviour of a real gas approaches that of an ideal gas when:

A. the temperature is low

B. the temperature is high

C. the volume is low

D. the pressure is high

E. the temperature and pressure are high

8. The rate of diffusion of hydrogen is about:

A. one-half that of helium

B. 1.4 times that of hydrogen

C. twice that of helium

D. four times that of helium

E. about 2.8 times that of helium

9. The critical temperature of H_2O is higher than that of O_2, because the H_2O molecule has:

A. fewer electrons than O_2

B. a dipole moment

C. two covanent bonds

D. a linear shape

E. a lower mass than O_2

10. Molecular solids are usually:

A. good electrical conductors

B. quite hard

C. quite brittle

D. volatile

E. high melting

11. In a covalent solid the units which occupy the lattice points are:

A. atoms

B. ions

C. molecules

D. electrons

E. positive ions only

12. An example of a covalent solid is:

A. hydrogen

B. sodium sulphate

C. diamond

D. copper

E. nitrogen

13. Dipole-dipole attractions exist in solid:

A. ammonia

B. carbon dioxide

C. hydrogen

D. sodium chloride

E. in none of these

14. In a molecular solid the forces holding together the atoms within the molecule are:

A. covalent bonds

B. ionic bonds

C. metallic bonds

D. electrostatic attraction

E. van der Waals forces

15. In a molecule of ammonium chloride there are:

A. dipole-dipole forces only

B. covalent bonds only

C. ionic bonds only

D. ionic and covalent bonds

E. van der Waals forces only

16. Which of the following statements are false:

A. Hydrogen has a very low normal boiling point, because its molecules are symmetrical and have very few electrons.

B. The molar volume of a liquid changes very little with changing pressure.

C. Water in an open thermos bottle will usually be colder than water in an open drinking glass.

D. The vapour pressure of chloroform ($CHCl_3$) at its normal boiling point is greater than that of water at its normal boiling point.

E. A much smaller fraction of the space occupied by a liquid is "empty" than is the case of a gas.

17. The volume occupied at S.T.P. by 1.00 g. of C_3H_8 gas is:

A. 0.508 litre

B. 0.988 litre

C. 1.01 litre

D. 22.4 litre

E. 2.24 litre

GRADE 12 PHYSICS

1. Which of the following is a vector quantity?

   A. momentum

   B. kinetic energy

   C. work

   D. heat

   E. temperature

2. A 10-kilogram weight is dropped from a height of 3 meters. Just before striking the ground the weight's kinetic energy will be about

   A. 3 joules

   B. 30 joules

   C. 300 joules

   D. 3000 joules

3. A number of the examples of the energy concept make use of "frictionless" systems. Why is this done?

   A. Most systems are frictionless.

   B. Total energy is not conserved when friction is present.

   C. Friction is not as meaningful a concept as energy.

   D. Friction is not present in outer space.

   E. It is often possible to get useful answers by ignoring friction.

4. Which one of the following is most nearly an "elastic" collision?

   A. two railway cars coupling

   B. an automobile collision

...... Continued

4. C. two billiard balls colliding

   D. an apple dropped on the ground

   E. a hammer hitting a nail into wood

5. When the displacement pattern of a transverse wave lies in a single plane, the wave is said to be

   A. reflected

   C. diffracted

   B. polarized

   D. refracted

6. Two steel balls collide elastically.

   A. Momentum is the same before and after the collision, but kinetic energy is not.

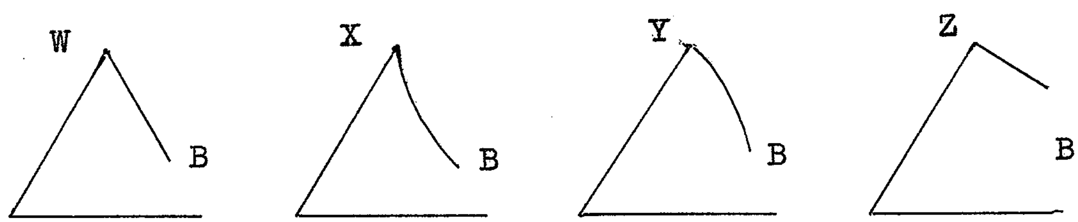
   B. Momentum and kinetic energy are the same before and after the collision.

   C. The temperature of both balls will increase.

   D. The balls will be permanently deformed.

   E. The balls will stick together.

7. A girl wants to slide down a playground slide so that she will have the greatest possible speed when she reaches the bottom (point B). Which of the following <u>frictionless</u> slides should she choose? (Points W, X, Y, Z are all two meters above the ground, and point B is 0.5 meters above the ground.)



7. A. slide W
B. slide X
C. slide Y
D. slide Z
E. The speed at B will be the same for each.

8. Leibniz's <u>vis viva</u> most closely resembles
A. potential energy
B. kinetic energy
C. heat
D. velocity
E. momentum

Questions 9 and 10 refer to the following statements:

A 0.1 kilogram snow ball strikes a 0.9 kilogram stationary skateboard and sticks to it. At the instant of impact, the snowball has a velocity of 18 meters per second in the horizontal direction. (Assume that the skateboard is on a perfectly horizontal stretch of ground and that it moves without friction.)

9. After collision, the skateboard and snowball move horizontally with a velocity of
A. 1.8 m/sec.
B. 2 m/sec.
C. 16.2 m/sec.
D. 18 m/sec.
E. 180 m/sec.

10. Kinetic energy is not conserved in the above collision because
A. the system is not closed.
B. the collision is not perfectly elastic.
C. momentum and energy cannot both be conserved in a collision.

.... Continued

10. D. the law of conservation of energy does not hold a frictionless system.

E. heat cannot flow from a cold object to a hot object.

11. An object is hung on a vertical spring and allowed to oscillate up and down. At any instant the object's total energy is

A. $KE + PE_{elastic} + PE_{gravitational}$

B. $KE + PE_{elastic}$

C. $KE + PE_{gravitational}$

D. $PE_{elastic} + PE_{gravitational}$

E. KE

12. A body's momentum is defined as the body's mass times its velocity. The mks unit of momentum is

A. kilogram-meter

B. kilogram-meter/sec.

C. kilogram-meter^2/sec^2

D. kilogram^2-meter/sec.

E. none of the above.

13. A satellite travels in a permanent, elliptical orbit around the earth. Which of the following statements are true of the satellite as it travels in its orbit?

I. Its total mechanical energy is the same at all times.

II. The gravitational PE increases as the satellite slows down.

III. Its kinetic energy is greatest when it is nearest the earth.

13. A. I only
    B. I and II
    C. I and III
    D. II and III
    E. I, II and III

14. <u>ALL</u> EXCEPT ONE of the following are in agreement with Goethe's nature-philosophy. Which one is the exception?
    A. The methods of mechanistic science (for example, mathematical analysis and experimentation) give the wrong idea of nature.
    B. Nature as it really is can be understood by direct observation.
    C. Gravity, electricity and magnetism are simply different manifestations of one basic quantity.
    D. One should search for the inner meaning of nature.
    E. Laws that are practical and quantitative can best describe nature.

15. The kinetic energy of an object is increased the most by doubling its
    A. mass
    B. temperature
    C. volume
    D. density
    E. speed

16. The first law of thermodynamics is a statement of
    A. the law of conservation of energy.
    B. the law of conservation of momentum.
    C. the law of conservation of mass.
    D. Newton's law of action and reaction.
    E. Galileo's law of motion.

17. Even though one may listen to a band from a considerable distance, the sound of the piccolo and that of the tuba do not get "out of step" with each other. This is evidence that in this situation sound waves

A. travel at the same speed for all frequencies.

B. are not polarized.

C. are longitudinal.

D. tend to be sinusoidal.

E. travel at a slower speed than light.

18. Two spheres of the same diameter, one of mass 5 kilograms and the other of mass 10 kilograms, are dropped at the same time from the top of a tower. When they are 1 meter above the ground, the two spheres have the same

A. momentum

B. kinetic energy

C. potential energy

D. total mechanical energy

E. acceleration

19. When a gas is held at a constant temperature, its molecules

A. have a certain constant average energy.

B. all have the same energy.

C. all have different energies that remain constant.

20. In 1620 Francis Bacon wrote: "There is nothing more true in nature than the twin propositions that 'nothing is produced from nothing' and 'nothing is reduced to nothing' . . . the sum total of matter remains

20. unchanged, without increase or diminuation." This statement implies which of the following basic scientific principles?

A. conservation of momentum

B. conservation of *vis viva*

C. conservation of mass

D. conservation of mechanical energy

E. conservation of charge

APPENDIX B

DIRECTIONS FOR SCORING ACCORDING TO THE SUBJECTIVE PROBABILITY TECHNIQUE

The items on this test have been designed to test both your knowledge of the subject matter and your confidence in that knowledge.

Read all responses for each item and distribute a total of five points per question in such a way as to ensure the maximum likelihood of achieving the best total score on the test.

One way of deciding on the allocation of weights for the various choices would be to consider each question as common stock valued at one dollar (at present). The correct answer could be represented as a stock which achieves a maximum value of five dollars. You are to bid a total of five dollars per item with the bids on each choice weighted according to your knowledge and confidence in the choices. Your return is determined by the amount you have bid on the correct response plus a small fraction of your bids on the other choices. Following are three examples of three possible procedures:

1. You are certain of the answer. You place $5 on it.
   - A. If you are right - you receive $5.
   - B. If you are wrong - you receive $0.

2. You have no idea of the answer. You place $1 on each choice. You receive $1 + $2 = $3.

3. You have some idea or have eliminated some choices. You place $2, $2, $1.

   If $2 on the right answer - you get $3.60

   If $1 on the right answer - you get $2.60

   If $0 on the right answer - you get $1.60

Any combinations of weights is acceptable, and returns are in proportion to the apparent knowledge indicated by your selection of weights. The purpose of the test is to diagnose your deficiencies so that they can be remedied. Please answer all items.

You may check your first choice on the first row of blanks and then indicate your distribution of weights on the second row of blanks.

Are there any questions?

APPENDIX C

TABLE XXVI

SUBJECTIVE PROBABILITY SCORES FOR THE VARIOUS COMBINATIONS OF WEIGHTINGS

Combination		de Finetti	Scale x 5
		$S = r_h + \frac{1}{2}(1 - \sum r^2)$	S x 5
5/0/0/0/0	$S_5 = 1$	$+ \frac{1}{2}(1 - 1) = 1$	5.0
	$S_0 = 0$	$+ \frac{1}{2}(1 - 1) = 0$	0.0
4/1/0/0/0	$S_4 = .8$	$+ \frac{1}{2}(1 - .68) = .96$	4.8
	$S_1 = .2$	$+ \frac{1}{2}(1 - .68) = .36$	1.8
	$S_0 = 0$	$+ \frac{1}{2}(1 - .68) = .16$	0.8
3/2/0/0/0	$S_3 = .6$	$+ \frac{1}{2}(1 - .52) = .84$	4.2
	$S_2 = .4$	$+ \frac{1}{2}(1 - .52) = .64$	3.2
	$S_0 = 0$	$+ \frac{1}{2}(1 - .52) = .24$	1.2
3/1/1/0/0	$S_3 = .6$	$+ \frac{1}{2}(1 - .4) = .95$	4.5
	$S_1 = .2$	$+ \frac{1}{2}(1 - .4) = .50$	2.5
	$S_0 = 0$	$+ \frac{1}{2}(1 - .4) = .30$	1.5
2/2/1/0/0	$S_2 = .4$	$+ \frac{1}{2}(1 - .36) = .72$	3.6
	$S_1 = .2$	$+ \frac{1}{2}(1 - .36) = .52$	2.6
	$S_0 = 0$	$+ \frac{1}{2}(1 - .36) = .32$	1.6
2/1/1/1/0	$S_2 = .4$	$+ \frac{1}{2}(1 - .28) = .76$	3.8
	$S_1 = .2$	$+ \frac{1}{2}(1 - .28) = .56$	2.8
	$S_0 = 0$	$+ \frac{1}{2}(1 - .28) = .36$	1.8
1/1/1/1/1	$S_1 = .2$	$+ \frac{1}{2}(1 - .2) = .60$	3.0

APPENDIX D

SAMPLE ANSWER SHEET

TEST.................................NAME....................

		A	B	C	D	E
1. M.C.		___	___	___	___	___
S.P.	___	___	___	___	___	___
2. M.C.		___	___	___	___	___
S.P.	___	___	___	___	___	___
3. M.C.		___	___	___	___	___
S.P.	___	___	___	___	___	___
4. M.C.		___	___	___	___	___
S.P.	___	___	___	___	___	___
5. M.C.		___	___	___	___	___
S.P.	___	___	___	___	___	___
6. M.C.		___	___	___	___	___
S.P.	___	___	___	___	___	___
7. M.C.		___	___	___	___	___
S.P.	___	___	___	___	___	___
8. M.C.		___	___	___	___	___
S.P.	___	___	___	___	___	___
9. M.C.		___	___	___	___	___
S.P.	___	___	___	___	___	___
10. M.C.		___	___	___	___	___
S.P.	___	___	___	___	___	___

APPENDIX E

DIAGNOSTIC TEST RESULTS

TABLE XXVII

GRADE X PHYSICS RESULTS

Student	M.C.	Systematic	S.P.Score	Rank	D	D^2
1	35	17	49	19	2	4
2	45	15	55	18	3	9
3	60	7	83	3	4	16
4	45	15	60	15	0	0
5	50	11	62	13.5	2.5	6.25
6	45	15	70	8	7	49
7	70	4	75	5	1	1
8	60	7	73	6	1	1
9	60	7	70	8	1	1
10	80	2	89	2	0	0
11	65	3	83	4	1	1
12	15	20	37	20	0	0
13	45	15	57	16	1	1
14	30	19	60	13.5	5.5	30.25
15	85	1	90	1	0	0
16	40	16	56	17	1	1
17	30	19	61	12	7	49
18	55	10	65	10	0	0
19	55	10	64	11	1	1
20	65	5	70	8	3	9
						179.5

$r_s = .865$

TABLE XXVIII

GRADE X GEOGRAPHY RESULTS

Student	M.C.	Systematic	S.P.Score	Rank	D	D^2
1	75	1	85	1	0	
2	70	3	75	5.5	2.5	6.25
3	70	3	75	5.5	2.5	6.25
4	70	3	80	2	1	1
5	60	6	74	8	2	4
6	60	6	75	5.5	.5	.25
7	65	8	68	12	4	16
8	60	6	68	12	6	36
9	55	9	75	5.5	3.5	12.25
10	50	12	67	15	3	9
11	50	12	59	20	8	64
12	50	12	67	15	3	9
13	50	12	68	12	0	
14	50	12	76	3	9	81
15	45	15	60	19	4	16
16	40	16	69	10	6	36
17	35	18	61	18		
18	35	18	65	16	2	4
19	30	20	71	9	11	121
20	30	20	46	22.5	2.5	6.25
21	25	21	62	17	4	16
22	0	21.5	56		.5	.25
23	0	21.5	46	22.5	1	1
						445.5

$r_s = .779$

TABLE XXIX

GRADE XI GEOGRAPHY RESULTS

Student	M.C.	Systematic	S.P.Score	Rank	D	D^2
1	65	1	79	1	0	
2	45	2	64	2	0	0
3	40	3.5	53	5	1.5	2.75
4	40	3.5	53	5	1.5	2.25
5	20	5	58	3	2	4
6	15	6.5	51	7	.5	.25
7	10	8	49	8	0	0
8	15	6.5	53	5	1.5	2.25
9	0	9	23	9	0	0
						15
	r_s	= .909				

TABLE XXX

GRADE XI CHEMISTRY RESULTS

Student	M.C.	Systematic	S.P.Score	Rank	D	D^2
1	100	3.5	100	3	.5	.25
2	100	3.5	100	3	.5	.25
3	100	3.5	100	3	.5	.25
4	100	3.5	100	3	.5	.25
5	100	3.5	100	3	.5	.25
6	100	3.5	100	6	7.5	6.25
7	90	7.5	85	7	.5	.25
8	90	7.5	85	8	.5	.25
9	85	9.5	80	9	.5	.25
10	85	9.5	80	11.5	2	4
11	80	11	76	13	2	4
12	75	12.5	70	15	2.5	6.25
13	70	15	65	10	5	25
14	75	12.5	70	16	3.5	12.25
15	70	15	65	18	3	9
16	70	15	65	14	1	1
17	65	17	60	19	2	4
18	65	18.5	55	11.5	7	49
19	65	18.5	55	20	1.5	2.25
20	60	20	50	21	1	1
21	45	21	30	22	1	1
22	35	22	20	17	5	25
						152
	r_s	= .906				

TABLE XXXI

GRADE XII CHEMISTRY RESULTS

Student	M.C.	Systematic	S.P.Score	Rank	D	D^2
1	100	1.5	99	1.5	0	
2	100	1.5	99	1.5	0	
3	82	3	88	3	0	
4	76	5	86	4	1	1
5	76	5	85	5	0	
6	76	5	80	6	1	1
7	70	7	78	7	0	
8	59	9	76	8	1	1
9	59	9	67	9	0	
10	59	9	67	9	0	
11	53	12	66	11	1	1
12	53	12	64	12	0	
13	53	12	58	15	3	9
14	41	15	63	13	2	4
15	41	15	62	14	1	1
16	41	15	55	16	1	1
17	36	17.5	53	17	.5	.25
18	36	17.5	45	19.5	2	4
19	29	19.5	48	18	1.5	2.25
20	29	19.5	41	21	2.5	6.25
21	24	22	45	19.5	2.5	6.25
22	16	21	40	22	1	1
23	18	23	39	23	0	
						39

$r_s = .981$

TABLE XXXII

GRADE XII PHYSICS RESULTS

Student	M.C.	Systematic	S.P.Score	Rank	D	D^2
1	45	7.5	51	7	15	.25
2	55	5	67	5	0	0
3	0	10	35	10	0	0
4	90	1	95	1	0	0
5	80	2	90	2	0	0
6	50	6	60	6	0	0
7	60	4	69	4	0	0
8	45	7.5	58	8	.5	.25
9	70	3	78	3	0	0
10	40	9	50	9	0	0
						.5

r_s = .997

APPENDIX F

COMPARISON OF DIAGNOSTIC TEST RESULTS WITH EXAMINATION RESULTS

TABLE XXXIII

GRADE X PHYSICS

Student	M.C.	X	S.P.	Y	EXAM	Z
1	35	-17	49	-17	60	- 1
2	45	- 7	55	-11	74	13
3	60	8	85	19	76	15
4	45	- 7	60	- 6	58	- 3
5	50	- 2	62	- 4	64	3
6	45	- 7	70	4	58	- 3
7	70	18	75	9	66	5
8	60	8	73	7	60	- 1
9	60	8	70	4	72	11
10	80	18	89	23	66	5
11	65	13	83	17	60	- 1
12	15	-37	37	-29	50	-11
13	45	- 7	57	- 9	55	- 6
14	30	-22	60	- 6	55	- 6
15	85	33	90	24	65	4
16	40	-12	56	-10	45	-16
17	30	-22	61	- 5	66	5
18	55	3	65	- 1	50	-11
19	55	3	64	- 2	60	- 1
20	65	13	70	4	64	3
Mean	52		66		61	

TABLE XXXIII (Continued)

Student	X^2	Y^2	Z^2	XZ	YZ
1	289	289	1	17	17
2	49	121	169	- 91	-143
3	64	289	225	120	285
4	49	36	9	21	18
5	4	16	9	- 6	- 12
6	49	16	9	21	- 12
7	324	81	25	90	45
8	64	49	1	- 8	- 7
9	64	16	121	88	44
10	324	524	25	90	115
11	169	289	1	- 13	- 17
12	1369	841	121	407	319
13	49	81	36	42	54
14	484	36	36	132	36
15	1089	576	16	132	96
16	144	100	256	192	160
17	484	25	25	-110	- 25
18	9	1	121	- 33	11
19	9	4	1	- 3	2
20	169	16	9	39	12
	5255	3411	1216	1037	1005

$r_{mc} = .411$ $r_{sp} = .416$

TABLE XXXIV

GRADE X GEOGRAPHY

Student	M.C.	X	S.P.	Y	EXAM		Z
1	75	28	85	18	59	12.5	- 3
2	70	23	75	8	55	18.5	- 7
3	70	23	75	8	70	3	
4	70	23	80	13	94	1	32
5	60	13	74	7	80	2	18
6	60	13	75	8	55	18.5	- 7
7	65	18	68	1	55	18.5	- 7
8	60	13	68	1	69	4	7
9	55	8	75	8	66	5	4
10	10	3	67	0	65	6	3
11	50	3	59	-12	55	18.5	- 7
12	50	3	67	0	64	7	2
13	50	3	68	1	55	18.5	- 7
14	50	3	76	9	60	10.5	- 2
15	45	- 2	60	- 7	63	8	1
16	40	- 7	69	2	57	14.5	- 5
17	35	-12	61	- 6	55	18.5	- 7
18	35	-12	65	- 2	62	9	0
19	30	-17	71	4	60	10.5	- 2
20	30	-17	46	-21	45	23	-17
21	25	-22	62	- 5	50	22	-12
22	0	-47	56	-11	59	12.5	- 3
23	0	-47	46	-21	57	14.5	- 5
Mean	47		67		62		

TABLE XXXIV (Continued)

Student	X^2	Y^2	Z^2	XZ	YZ
1	784	324	9	- 84	- 54
2	529	64	49	-161	- 56
3	529	64	64	184	64
4	529	169	1024	736	416
5	169	49	324	234	112
6	169	64	49	- 91	- 56
7	324	1	49	-126	- 7
8	169	1	49	91	7
9	64	64	16	32	32
10	9	0	9	9	0
11	9	144	49	- 21	84
12	9	0	4	6	0
13	9	1	49	- 21	- 7
14	9	81	4	- 6	- 18
15	4	49	1	- 2	- 7
16	49	4	25	35	- 10
17	144	36	49	84	13
18	144	4	0	0	0
19	289	16	4	34	- 8
20	289	441	289	289	357
21	484	25	144	264	60
22	2209	121	9	141	33
23	2209	441	25	235	105
	9142	2168	2194	2374	1283
				-512	-223
				1862	1060

TABLE XXXV

GRADE XI GEOGRAPHY

Student	M.C.	X	S.P.	Y	EXAM	Z
1	65	37	79	25	62	7
2	45	17	64	10	55	0
3	40	12	53	- 1	58	3
4	40	12	53	- 1	56	1
5	20	- 8	58	4	60	5
6	15	-13	51	- 3	51	0
7	10	-18	49	5	50	- 5
8	15	-13	53	- 1	57	2
9	0	-28	29	-25	40	-15
Mean	28		54		55	

Student	X^2	Y^2	Z^2	XZ	YZ
1	1369	625	49	259	175
2	289	100	0	0	0
3	144	1	9	36	- 3
4	144	1	1	12	- 1
5	64	16	25	-40	20
6	169	9	0	0	0
7	324	25	25	90	-25
8	169	1	4	-26	- 2
9	784	625	225	420	375
	3476	1403	338	741	499

$r_{mc} = .686$ $r_{sp} = .726$

TABLE XXXVI

GRADE XI CHEMISTRY

Student	M.C.	X	S.P.	Y	EXAM		Z
1	100	27	100	17	78	4.5	17
2	100	27	100	17	87	1.5	26
3	100	27	100	17	60	9	- 1
4	100	27	100	17	70	8	9
5	100	27	100	17	85	3	24
6	100	27	99	16	58	12.5	- 3
7	85	12	91	8	87	1.5	26
8	85	12	90	7	75	6	14
9	80	7	85	2	50	18	-11
10	80	7	83	0	45	21.5	-16
11	75	2	80	- 3	71	7	10
12	70	- 3	77	- 6	78	4.5	17
13	65	- 8	84	1	50	18	-11
14	70	- 3	75	- 8	52	14	- 9
15	65	- 8	70	-13	50	18	-11
16	65	- 8	79	- 4	50	18	-11
17	60	-13	73	-10	50	18	-11
18	55	-18	83	0	59	10.5	- 2
19	55	-18	69	-14	58	12.5	- 3
20	50	-23	61	-22	59	10.5	- 2
21	30	-43	45	-38	51	15	-10
22	25	-53	74	- 9	45	21.5	-16
Mean	73		83		61		

TABLE XXXVI (Continued)

Student	X^2	Y^2	Z^2	XZ	YZ
1	729	289	289	469	289
2	729	289	676	702	442
3	729	289	1	- 27	- 17
4	729	289	81	243	153
5	729	289	576	628	408
6	729	256	9	- 81	- 48
7	144	64	676	312	208
8	144	49	196	168	98
9	49	4	121	- 77	- 22
10	49	0	256	-112	0
11	4	9	100	20	- 30
12	9	36	289	- 51	-102
13	64	1	121	88	- 11
14	9	64	81	27	72
15	64	169	121	88	143
16	64	16	121	88	44
17	169	100	121	143	110
18	324	0	4	36	0
19	324	196	9	54	42
20	529	484	4	46	44
21	1849	1444	100	430	380
22	2809	81	256	658	144
	10978	4418	4188	4220	2577
				348	190
				3872	2387

$r_{sp} = .550$ $\quad r_{mc} = .556$

TABLE XXXVII

GRADE XII CHEMISTRY

Student	M.C.	X	S.P.	Y	EXAM	Z
1	100	46	99	34	86	16
2	100	46	99	34	94	24
3	82	18	88	23	90	20
4	76	22	86	21	78	8
5	76	22	85	20	84	14
6	76	22	80	15	72	2
7	70	16	78	13	85	15
8	59	5	76	11	75	5
9	59	5	67	2	75	5
10	59	5	67	2	60	-10
11	53	- 1	66	1	63	- 7
12	53	- 1	64	- 1	80	10
13	53	- 1	58	7	55	-15
14	41	-13	63	- 2	67	- 3
15	41	-13	62	- 3	51	-19
16	41	-13	55	-10	70	0
17	36	-18	53	-12	70	0
18	36	-18	45	-20	50	-20
19	29	-25	48	-18	80	10
20	29	-25	41	-24	51	-19
21	24	-30	45	-20	67	- 3
22	26	-28	40	-25	59	-11
23	18	-36	39	-26	52	-18
Mean	54		65		70	

TABLE XXXVII (Continued)

Student	X^2	Y^2	Z^2	XZ	YZ
1	2116	1156	256	690	544
2	2116	1156	576	1104	816
3	324	529	400	360	460
4	484	441	64	176	168
5	484	400	196	308	280
6	484	225	4	44	30
7	256	169	225	240	195
8	25	121	25	25	55
9	25	4	25	25	10
10	25	4	100	- 50	- 20
11	1	1	49	7	- 7
12	1	1	100	10	- 10
13	1	49	225	15	-105
14	169	4	9	39	6
15	169	9	361	247	57
16	169	100	0	0	0
17	324	144	0	0	0
18	324	400	400	190	400
19	625	324	100	-250	-180
20	625	576	361	475	456
21	900	400	9	90	60
22	784	625	121	308	275
23	1296	676	324	648	468
	11727	7514	3930	4701	4270
					325
					4148

r_{sp} = .695 r_{mc} = .701

TABLE XXXVIII

GRADE XII PHYSICS

Student	M.C.	X	S.P.	Y	EXAM		Z
1	45	-10	51	-14	70	6	5
2	55	0	67	2	75	3	10
3	0	-40	35	-30	55	7	-10
4	90	35	95	30	74	4	9
5	80	25	90	25	84	2	19
6	50	- 5	60	- 5	50	8	-15
7	60	5	69	4	44	9	-21
8	45	-10	58	- 7	73	5	8
9	70	15	78	13	85	1	20
10	40	-15	50	-15	40	10	-25
Mean	55		65		65		

Student	X^2	Y^2	Z^2	XZ	YZ
1	100	196	25	- 50	- 70
2	0	4	100	0	20
3	1600	900	100	400	300
4	1225	900	81	325	270
5	625	625	361	475	475
6	25	25	225	75	75
7	25	16	441	-105	- 84
8	100	49	64	- 80	- 56
9	225	169	400	300	260
10	225	225	625	675	1000
	4150	3009	2422	2015	2190

r_{sp} = .635 $\quad$ r_{mc} = .884

)HN'S–RAVENSCOURT SCHOOL

SOUTH DRIVE, FORT GARRY, WINNIPEG 19, MANITOBA

F. M. OLSEN, B.Sc., B.Ed
Director of Admissions
Residen