1939

Achievement tests in modern languages

Leader, Max

Boston University

http://hdl.handle.net/2144/18087

Boston University
BOSTON UNIVERSITY
GRADUATE SCHOOL
Thesis
ACHIEVEMENT TESTS in MODERN LANGUAGES
by
Max Leader
(B.S., Boston University, 1931)
submitted in partial fulfilment of the
requirements for the degree of
Master of Arts
1939
APPROVED

by

First Reader Samuel M. Wayman
Professor of French

Second Reader André Célière
Professor of French
Contents

Volume I

1. Present Day Conditions in Measurement 1
2. Traditional Examinations and Marks 7
3. The New-Type Examination 14
4. Conclusion 31
5. Informal Objective Test 36
6. Bibliography 41

Volume II Appendix

Exhibit of Objective Tests
A. The Standardized Type
B. Non-Standardized Tests
C. Progress Tests to Accompany Text Book
Achievement Tests in Modern Languages

I. Present Conditions in Measurement

Education apparently cannot be efficiently carried on without some form of measurement. Tests and measurements have been used from time immemorial. History tells us that examinations were prevalent as far back as the eras of Greek and Roman dominance. The Chinese, too, before the Greeks and Romans, had a very ancient and intricate system of examinations. In fact, specific testing of what has been learned has had its place in practically every scheme of learning of which we have record.

In this day of measuring with an educational yardstick, the uncertainty which prevails in the field of modern languages makes them very vulnerable to attack. This may be due to the fact that until a few years ago the elementary grades were considered far more critical in the life of the pupil and less attention was paid to the problems of the high school student. With the rapid growth in the enrollment in secondary schools, we have come to realize that the twelve year course is the recognized public school education of the average child. With increasing groups we must be prepared to measure and classify our students so that they may receive maximum benefit. The prevailing system of measurement in most cases does not provide for individual differences or abilities.

In the teaching of modern languages, there has been little agreement either as to objectives or as to the content of the
high school course. However, the type of examination and system of measurement has changed very little, if at all. Both are dominated by the colleges and universities. For example, Harvard began to require French or German for entrance in 1875. Down to the beginning of the World War few changes were made in the kind of examinations set by schools and colleges. The examination paper was modeled after the lesson plan characteristic of the synthetic type of grammar in use during the period. The brief examination in French required for admission to Harvard in 1876 consisted of first, some sixteen lines of prose to be translated into English; second, grammar questions requiring the conjugation and principal parts of verbs; and third, a few detached sentences to be turned into French. In 1899, the Committee of Twelve incorporated in its report specimen papers for the three grades, which were constructed under the same general model. Shortly thereafter the College Entrance Examination Board and the colleges put these into practice. Improvements, if any, in the nearly first fifteen years of the Board's existence were more concerned with quality than with form. The elementary French paper for 1911 introduced a novelty in the form of a simple question on pronunciation, such as, "What are the four nasal vowels?" Two years later a question of completion was introduced. From then on, however, tests of language skill have been made more or less consistently by means of functional grammar questions involving completion and mutation of sentence material, by
general questions of the type "Depuis quand étudiez-vous le français," or by questions on the passage to be translated.

Our methods of teaching and our textbooks have undergone great changes in the past few years, yet our system of examinations and methods of measurement are still committed to the same fixed and static program. Let us consider a typical comprehensive high school. The teaching of and achievement in the modern languages are still in a state of chaos. Teachers are finding it extremely difficult to justify their positions in the light of actual accomplishments. William R. Price has said, "There is no problem of secondary education that is causing more discussion than the place of foreign languages in the high school. It would almost seem that the only friends of foreign languages are the teachers of these languages. Sometimes, also, it would seem that the worst foe of these languages are their friends, the teachers." *

Pupils in the eighth and ninth grades of the junior high school elect a modern language and receive credit for promotion to the high school. Regardless of ability or achievement, these same pupils are required to repeat the first year of the language in high school. Although repetition is necessary in the study of a foreign language, such a procedure does more harm than good; makes teachers guilty of waste and inefficiency.

* "Foreign Language Teaching in American High Schools"
High School students who elect foreign languages divide themselves into three groups, namely, the college preparatory sections, the commercial sections, and the general sections. The college preparatory group has a definite goal to reach and so it is they are willing and able to work. Occasionally, there are students who are misplaced or who cannot keep up with the rest of the class but they soon drop out or change to the general course. Unfortunately, this college preparatory group is the smallest. As for the commercial group much need not be said because those students are so overburdened with required subjects in their particular course that there is little time in their program for languages. Furthermore they avoid languages because they do not wish to put sufficient time in study. The Third group presents the greatest problem. In a comprehensive high school, the majority of students are likely to elect the general course. With very few exceptions, they are free to elect any subject in the high school curriculum. Most elect a foreign language simply as a means to an end, so many credits toward graduation. They receive credit for attending classes regularly and for doing a minimum amount of work, not for actual achievement. In the first year of the modern language, a minimum number of lessons in a standard grammar are studied and very little reading, perhaps fifty pages, have been completed. For this work most of the students receive a passing grade. The number of lessons completed does not necessarily mean that they have been mastered, for in the second year it takes approximately two thirds of the school year to review the first year's work.
This leaves very little time to complete the minimum essentials of a two year grammar course. The reading, however, consists of a little more than the first year.

Low marks or failing marks cause pupils to discontinue the study of a modern language and leave the teacher open to criticism from the principal, higher school authorities or parent teachers' organizations. It must be admitted that enrollment in the modern languages is decreasing annually in almost all states. With the exception of the college preparatory students, modern languages are elective. In order to keep enrollment figures as high as possible, teachers tell their students that they will not receive credit for one year of a language toward the high school diploma. On the one hand, teachers promise good grades to their pupils if they continue the study of modern language and on the other, they are threatened with low marks and even failure unless they do.

The situation in our colleges is far different from that in the secondary schools. Most colleges that do not require entrance examinations or that admit students on certificate, count two years of high school foreign language credit as equivalent to one of their own. No credit for entrance is granted for less than two years of a foreign language. The one year high school student must begin over again in the language work and register for the first semester of elementary language. Colleges are universal in requiring two years of a language for entrance. The curriculum of the secondary school
and the content of its courses are made up in accordance with college requirements. But all secondary schools do not have uniform objectives and there are differences in textbooks, in the amounts of material covered and in the quality of instruction. This causes a wide variety in the quality of the students' preparation. In the Chelsea Senior High School, not a single student who has taken the College Entrance Board Examinations in French has failed in the past twenty-five years. The training in grammar fundamentals is incomparable. The weakness there is in pronunciation and diction.

Usually colleges do very little, if anything, to find out for themselves and for the students the amount and quality of this previous preparation. Courses are offered as if all students are equally prepared. Within a week or two, the students are either well established or on the way to failure.

Entrance examinations provide one means of checking up on previous preparation: but they must be uniform in difficulty and based on similar objectives from year to year to be comparable. If graded by more than one individual, there will be considerable error due to differences of judgment on the part of the examiners.
II Traditional Examinations and Marks

At the present time, written examinations occupy a prominent place in the work of the secondary schools and colleges. Students accept them as a matter of course, and the accuracy of the marks placed upon examination papers is not seriously questioned. There are many high school teachers and college professors who feel that their system gives very satisfactory results in their own personal use of it. Furthermore, these conservatives resent any proposal to change a system of education which they credit with producing the educated men of today.

Countless experiments have shown the unreliability of these written examinations. Starch and Elliot* had facsimile copies of a final examination paper in first year high school English graded by 142 English teachers in as many high schools. The scores ranged from 50% to 98%.

A final examination paper in American History was graded by 70 history teachers; one teacher assigned it a grade of 43 and another 90; a dozen teachers rated it as 80 or above, and another scored it as below 55.

It might be objected that these are not fair samples of the system because of the peculiar difficulties of measuring

English, history, foreign language, and similar subjects with necessarily indefinite content. These objectors would agree that mathematics is one of the most definite and objective subjects in our curriculum, and that a trial of mathematicians or teachers of mathematics would be fairer. Another experiment by Starch and Elliot* was a final examination paper in geometry scored by 114 mathematics teachers; as in the case of the English and history papers, it was marked on the basis of 100 as perfect. One teacher marked it as low as 28, and two marked it as high as 92. A dozen teachers marked it as 53 or below, and fourteen marked it as 83 or above. Obviously there is no need for comment on variations so great as these. Similar variations would be discovered in papers in other subject matters scored by teachers in the respective departments of the same school.

Now it is only fair to consider what a written examination constitutes and how marks are derived. In present practice, the written examination usually consists of from five to ten questions on as many important points of the course in which the examination is given. The questions are chosen by the individual instructor; they are worded by him; their relative value is determined by him as many times as there are papers scored; and the answers to the questions are evaluated by him.

subjectively. The reading of a poor examination paper may have an adverse effect on a succeeding good paper and vice versa. It is possible that the reader is in a hurry and the students may gain or lose by his haste. The physical condition of the reader may also have an effect on the mark of an examination paper. This method is no better than that of personal judgment. Many times the written examination is the sole measure, as, when the student is unknown to the scorer. The written examination is supposed to show the student's "power of organization." What this "power" is, is left to the "good sense" of the individual instructor. In practice, it would appear that an influential feature of the written examination is sheer mass of reproduction coupled with any manner of evidence of the student's shrewdness in guessing the intent of the questions. In examinations, many college professors are much pleased to read, in answers to questions, the very same words which they had used in lectures during the semester. From such instruments of marking reliable measurements can hardly be expected. It is also true that there is no organization in the new-type tests.

In the last analysis grades are whatever the individual instructor makes them and can be reduced to the personal judgment of the instructor at the moment of recording the grade. At the present stage of civilization traditional methods cannot be accepted as adequate. The mere fact of being an instructor does not necessarily mean that an individual is
automatically qualified to give accurate, reliable, and meaningful marks according to the modern conceptions of such terms.

The results of most studies on the value of written examinations criticize them for their inaccurate measures of school achievement. Monroe*, in summarizing criticisms of written examinations by various writers noted five reasons for this:

1. The marking of papers is highly subjective.
2. The questions of ordinary examinations are usually not equal in difficulty and weighting by teachers is highly subjective.
3. Frequently the content of written examinations is not in agreement with recognized educational objectives.
4. Written examinations are usually so short that they do not offer an adequate opportunity for a student to demonstrate his ability.
5. Distinction between "scores" on examinations and "grades" which interpret the scores with reference to norms or standards is not recognized. Such norms are very indefinite and highly subjective. This means that when a teacher writes a set of questions, marks them in numerical values or letters, and then subjects them to his scale of grades, he claims that his questions are of just the right difficulty, length and comprehensiveness that the average student should make a mark

midway in the class list, etc. The teacher has thus set the norm beforehand. If half the class should get below the passing mark, he might wonder what was the matter with his class, but probably would not realize that his questions were too difficult and that the passing mark should be much lower for that examination. The examinations of many teachers consist of "catch" questions with which to trap their pupils. Other teachers make their examinations too long to be completed within the allotted time. Each paper is marked according to the amount completed and the quality of it.

Teachers of foreign languages, to test knowledge of grammar, usually write a set of English sentences to be translated into French or some other foreign language. The teachers are the ones who must decide whether the sentences are of the right difficulty. If the sentences are easy a high set of marks may be obtained; if they are difficult, all the marks may be low. If these marks are called grades directly, the teacher might arbitrarily set the norms of the test at 92-100 for A; 85-91, for B; and so on, or whatever values they are accustomed to use. It is quite apparent that it is very difficult for teachers to write exercises of just the right difficulty to suit such established norms.

Classes differ in ability and teachers' methods of marking differ. Some teachers mark every mistake made in the translation of a set of sentences. Sometimes they set the mark by the "redness" of the paper after their red pencil has done its duty. Sometimes they count the mistakes and make a
descending scale of marks to fit the ascending number of mistakes. Some teachers select certain words and expressions in the exercises to be looked for, assuming that the rest will be right or wrong in proportion that these are right or wrong. Some teachers consider major or minor errors. Undoubtedly, there are many who merely read the translation and mark it "by the way it sounds." There are even a few teachers who do not correct the papers but simply assign a mark for the course on the basis of the attitude and behavior of the pupil. So many methods of marking, even if the questions are the same for all classes, as in a combined final examination, can not help causing many variations in the marks set by teachers. The difficulty is greater when the teachers set their own weights to the questions, as Question I, 20%; Question II, 40%, etc.

Monroe*, in his study, further found that of 145 high school teachers of modern languages only 15 wrote out the correct answers before starting to mark papers; 74 read all the questions on one paper before proceeding to the next paper; 23 read one question throughout the set of papers before proceeding to the next question; and 33 read the papers as a whole without marking questions separately.

Thus, we can clearly understand that the tests which are

most commonly given today are used to determine whether a pupil passes or fails, whether he is to be promoted or not. To summarize it has the following weaknesses:

I It contains too few questions to measure satisfactorily achievement in the subject tested.

II It measures work only in one grade, regardless of either a pupil's lack of mastery of the work of a preceding grade or his knowledge of the work of a higher grade.

III The scoring is subjective--teachers disagree seriously in the mark they assign the same answer paper.

IV Too much and too difficult written expression is demanded of pupils. This leads to subjective scoring and to the pupils often not being able to show how much he knows in the subject.

V The standard is arbitrary; it disregards the difficulty of the test and is, therefore, unfair.

VI The method of stating the test scores does not permit them to be used for the purposes for which test results should be used--to determine educational needs of pupils and instructional needs of teachers, and to provide a sound basis for supervision.
III The New-Type Examination

During the past few years examinations in modern languages have been undergoing drastic changes. There is at present a widespread interest in what is commonly called the "new-examination" or the "new-type test." There are two essential characteristics of these tests. In the first place, they call for very brief pupil responses, ordinarily single words, numbers or marks. Second, they can be scored with a high degree of objectivity; in other words, practically no doubt as to the correctness of any responses which pupils may give.

Any good examination must measure the ability and accomplishment of pupils. It must offer remedial help, especially for the unsatisfactory ones. It must measure teacher efficiency. It must provide opportunities for learning. It must motivate study and determine standards or goals of achievement.

A good examination must also be objective so that competent persons do not disagree as to what answers are correct and what are incorrect. The test must measure whatever it sets out to measure accurately. It should yield the same results when given to the same pupils two or more times, provided nothing has intervened which affects the knowledge or ability being tested. A test should be valid in that it measures what it claims to measure. An examination should be easy for the teacher to give and to score, for the pupils to take, and should be economical of the time of both. It should ordinarily
be easy enough that all pupils can score something, and none, or very few, make perfect scores. Questions and exercises should be clearly and definitely stated, and the general directions accompanying a test should be such that no one will misunderstand them.

The modern foreign languages lend themselves readily to the adaptation of new-type tests but teachers are hesitant to change from the existing system. In 1927, Ruch & Stoddard* made the following statement in their book: "Measurement in French and Spanish is in a state of flux. The test materials now available for the high school teacher are few in number and rather narrow in scope." Very few standard tests were then available. Today there are many foreign language tests which have been standardized and which have already proved their value as an aid to learning and teaching.

The "new-type" or objective tests as they will hereafter be called are classified into two main groups with subdivisions under each as follows:

I Recall types:

1. Simple recall
2. Completion exercises

II Recognition types:

1. Multiple response
2. Matching exercises
3. True-False

A single-answer or recall test is one composed of direct questions or their equivalents to each of which the answer is a single expression, usually a word, concerning the correctness of which there is no doubt. Although this is classed as one of the types of the new examination, it is not at all new. Practically all teachers have employed it in one form or another. When a foreign word is given and the English equivalent asked or vice versa, when a particular form of a certain word is called for, when a question is asked as to the case or tense of a word, the exercise falls under this classification. To illustrate consider these examples:

Directions: Replace the English words by the French words observing all rules of agreement and position:

1 Il a acheté une (gold) montre.
2 Où sont vos (kid) gants?

Write in the blank spaces antonyms of the following words:

1 jeune ------- 3 noir ------- 5 vite -------
2 beau ------- 4 chaud ------- 6 ouvrir -------

Completion tests are in many ways similar to the single-answer tests, differing chiefly in form. For example: Fill in the blank spaces with correct French word for the English word in parentheses:
1 (which) Avez-vous apporté le livre ____ est dans le salon?

2 (her) Je ____ ai dit que ____ soeur avait invité ____ et moi à aller avec ____ à ____ maison.

Probably the best and most frequently useful variety of new-type tests is the multiple answer. Such tests may be constructed to cover almost every type of material dealt with. In them the pupil is expected to copy the number of the correct answer on the short blank line provided on the paper, thus bringing all the answers into a straight column. Sometimes the pupil places a checkmark beside the correct French translation of several given. For example--Select the correct French translation of the English sentence and place a check in the space after it on the right.

1 My aunt arrived yesterday.
   Ma tante a arrivé hier.
   Ma tante arrivait hier.
   Ma tante est arrivée hier.
   Ma tante est arrivé hier.

This form may also be used in vocabulary tests in which the French word is given followed by 5 English translations, the correct one is to be selected and the number placed at the right.

Matching exercises are good but in foreign languages their use is somewhat limited because if a pupil knows most of the items he can determine the rest by the process of elimination. However, to overcome this all that has to be done is to make
one column longer than the other.

Match the French idioms with their English equivalents:

_____ près de 1 only
_____ il faut 2 never
_____ chez 3 to need
_____ tout a coup 4 not at all
_____ d'abord 5 near
_____ avoir besoin de 6 suddenly
_____ pas du tout 7 inorder to
_____ tout le monde 8 many people
_____ ne---jamais 9 it is cold
_____ il fait froid 10 at the home of

11 everybody
12 it is necessary
13 at first

True-false tests are less satisfactory than the other types. They may be used but not too often. If they are, the incorrect statements should not be too absurd.

The unstandardized objective test is a very good check on the pupils as well as the teacher. College professor insist on a knowledge of certain fundamentals. Through the use of these tests it is possible to check pupils more often, more easily and more efficiently. Errors are easily checked and remedial work assigned to the pupil.

There has not been much said anywhere about the measurement of reading either in class or outside reading. There is
a definite amount prescribed but no way of measuring it. The present day examinations of what has been read was considered satisfactory twenty-five years ago but not today. The questions depend on the attitude, health and whim of the examiner. Through the use of objective tests, which require very short answers, more questions can be asked and a wider sampling of the reading obtained. Professor Arsène Croteau of Connecticut State College reports success with the use of these tests and prefers them to the essay test.

At the State University of Iowa there is in use a very elaborate system of objective tests. The fundamental principle of the courses in the first two years of French is to emphasize individual initiative and responsibility by enabling all students to proceed in their study of the subject at the rate at which they seem best capable. This is known as the "Individualized Study Plan." The main features of this Plan are the following:

1. The division of each semester's work into projects, subject to study and checking by the student himself.

2. The classroom recitation on lesson assignments is eliminated. The recitation period consists of (1) oral practice in pronunciation; (2) supervised study and drill; (3) student self testing of progress in the projects; (4) personal conferences on difficult points and errors; (5) personal advice on study methods, reading materials, etc.

3. Achievement is measured through Departmental Examinations dealing with silent reading ability, vocabulary,
pronunciation, recognition grammar, content of supplementary reading texts, etc. Silent reading (comprehension) tests are not based on the content of material read in the text-books; they seek to measure the accuracy of a student's reading gained through practice in reading given material. If a student fails to pass a departmental examination he is expected to review the material as directed by his instructor and then apply for a re-examination.

4 To facilitate individual progress students will be sectioned at the beginning of the year on the basis of aptitude and placement examinations. The early sectioning casts no reflection upon any student because, due to the fact that each individual works according to his own initiative and ability, he may step out far in advance of the members of his class. In the event that he does this he will be transferred to a section where he will find students working approximately at his own rate and level.

Students are also encouraged to discuss any problems pertaining to their study of French with instructors during office hours. This is meant for all students from those who study foreign language as a requirement for the degree to the superior students.

The most significant result under the Iowa Plan is that credit is not given for time spent but for mastery of the required essentials of the course.

The earliest of the standardized tests to appear were the Handschin Modern Language Tests in 1920, in Comprehension and
Grammar and Reading. These tests are too short and the results prove very meagre. It seems to me that their only value is that they are a beginning in the new movement. These tests and others which will be described will be found in the appendix of this paper.

The following year the Henmon French Tests were published. There are four in number and consist of a vocabulary test and sentence test on the opposite sides of the same sheet. The vocabulary test consists of 60 words drawn from 448 words drawn from 12 widely used texts. The sentence tests consists of 12 sentences which contain words in the list of 448. Each word in the vocabulary test and each sentence has a numerical value. No time limit is set for the test but it is usually about 20 minutes. These tests are a slight improvement over the previous ones and also give a rough measure of class standing.

The Iowa Placement Examination has been revised so that it is now known as the New Series, F.T., Form X. It can be given after one-half year of French in high school or college, and covers a range of talent sufficient to measure four years of high school French. The test is divided into four parts. Part I consists of 40 French words for which the number of the English equivalent is to be placed on the same line at the extreme right of the page. Six minutes is allowed for this. Part II is a test of French grammar consisting of 40 items. The correct response is to be selected out of 4 and the number placed at the right of the page. The time limit is 14 minutes.
Part III contains 30 sentences in which the verb has been omitted. Again the correct response must be selected in 10 minutes this time. Part IV measures reading comprehension. Three paragraphs are given in increasing difficulty. Then statements based on the paragraph are marked true or false. There are 30 questions and these should be answered in 15 minutes. The test can be administered in an hour of class.

This test can be used in high schools which desire to measure aptitude and training in French and Spanish. These examinations are educational achievement tests standardized on the basis of the performance of first year college students. By using such a test college instructors are able to evaluate the French or Spanish knowledge possessed by individual students and of the language instruction in various high schools.

The Iowa Placement Examinations require no previous training on the part of the individual who administers them. They are easily scored by a key and scoring is highly objective.

The Columbia Research Bureau Tests in French, German and Spanish are the results of four years of work with experimental editions given to students in different parts of the country. The purpose of these tests was to give colleges and high schools more reliable, valid, and comparable measures of achievement than is given by the old-type examination.

The test appears in two forms and consists of three parts, vocabulary, comprehension and grammar. The vocabulary test consists of 100 words after each of which 5 English words
are given one of which is correct. The vocabularies of the most common grammars and composition books were compared for the final selection of the words. The comprehension test consists of 75 true-false statements. According to the authors of the test the statements are clearly within the understanding of all students who have sufficient intelligence to be in foreign language classes, so that the correct marking of each statement depends upon the student's ability to read French sentences couched in words to which all students have been about equally exposed. Part III is a grammar test of the completion type arranged in order of difficulty. The grammar elements were selected with a view to giving as representative a sampling of the common essentials as possible.

In 1925, 30,000 students in New York state high schools were given this test. From the scores norms were established and from the scores letter grades could be assigned. The test is very comprehensive. It requires 90 minutes to administer this test which is almost twice the normal class period. This may be considered the only fault of the test but it may be eliminated by giving the test on two successive days.

The American Council French Grammar test is of the selection type. It consists of 50 English sentences, each one being followed by four French translations, only one of which is correct. The student is to select the right translation under each item, checking the same in the parentheses opposite. The items for the test were selected after examination of some thirty grammars and composition books, the
College Entrance Board Examinations, the New York Regents Examinations, and other objective tests. The test requires 22 minutes for administration and measures achievement in French grammar only. It is a good test for college freshmen to determine their previous grammar knowledge and whether they should be allowed to take courses in French literature directly or whether they should be required to review principles of grammar. It can be used successfully in high schools to measure the progress of the pupils.

The American Council Alpha and Beta French Tests are similar to the Columbia Research Bureau Test so that space is not necessary to describe them.

The Columbia Research Bureau Aural French Test has been designed to measure the extent to which students understand spoken French. The student of French should be able 1 to follow directions that are given orally, 2 to understand an explanation, 3 to understand a question, 4 to understand the relation of facts, 5 to understand a continuous lecture or speech, and 6 to understand dictated sentences and to copy them. The test includes those functions which are stressed most in class work and in forms that can readily be measured.

Part I of the test consists of three pictures. The examiner reads a series of statements concerning the pictures which are before the student in his examination booklet. The student indicates whether each statement is true or false by
writing in numbered parentheses + (for true) and 0 (for false.)

Part II consists of 45 questions to be answered + (for yes) and 0 (for no) in numbered parentheses in his booklet. The first ten questions are read very slowly, one every 15 seconds; the next ten, slowly, one every 10 seconds: the following fifteen, one every 8 seconds; the last ten, one every 5 seconds.

Part III is a paragraph meaning test. The examiner reads a paragraph in French. Then the student turns to Part III in his booklet where he finds three sets of statements in English with four answers. The correct answer is to be selected and indicated. There are fifteen such paragraphs which do not have to be read in order. The examiner may skip around and announce the number of the paragraph after it is read. This eliminates the possibility of the student answering the questions while the paragraph is being read.

Part IV is a dictation exercise containing fifteen sentences of graded difficulty. The examiner reads each sentence once in its entirety and then in separate word groups. The text of the dictation has been divided into units for scoring purposes. These units are separated by vertical lines. The scoring key gives directions as to how much is to be deducted for errors in the units.

The test is long but the answers are so brief that it may be administered in fifty minutes. It is the first standardized aural test. It cannot be given by just any French teacher, for pronunciation and diction are the keynotes to the success
of it. Most French teachers in high schools today are lacking in these requirements. In the first two years of French, teachers assign grades on grammar achievement. In the literature courses the grade depends on the student's knowledge of the literature of the period. This test therefore is a very good one because it emphasizes one aspect of French which has been neglected for many years.

The Cooperative Test Service was established in 1930 to render a significant type of educational service to secondary schools and colleges. Its primary purpose is to make available the best instruments for the measurement of educational achievement that present knowledge and skill will permit.

The Cooperative foreign language tests rely largely on the pioneering work of the Modern Foreign Language Study. Each examination in the foreign languages is designed to test ability to read the language and those phases of the ability to write the language which lend themselves to objective testing. The tests are so constructed as to emphasize functional knowledge rather than mere memorization of forms or values. Each test consists of three parts: reading, vocabulary, and grammar. The reading part is designed to give a direct measure of the ability to understand the written language. The second part measures the extent of the student's vocabulary. The last part covers grammar knowledge from the functional point of view, dealing with matters of inflection, syntax, word order, etc.
The Cooperative language tests are published at two levels of difficulty, one suited to students with less than four semesters of high school study, or one year of college training, and the other suited to more advanced students. In addition, the tests at each level have been prepared in two lengths, a 90-minute edition and a shorter 40-minute form.

An excellent feature of the Cooperative tests which the others have not is to make available each year a new edition of each test. This policy of annual editions permits continuous adaptation of the tests to changing curriculum practices and makes possible the immediate application of the most recent developments in testing techniques. Furthermore, it makes continually available to the test users a supply of fresh test materials. The tests are also provided with norms which will render the results highly comparable from year to year.

All the standardized tests offer a common scale on which achievement at all levels may be measured. They measure growth of individuals as well as groups. The tests may be used to measure a student's achievement relative to his own group and relative to larger groups; to measure gains in achievement over a period of time; to provide accurate placement in classes according to actual achievement rather than according to the amount of time spent in the classroom; for determining levels of achievement in meeting special requirements; and for various other purposes for which objective measures are useful.

The State University of Iowa is doing the most progressive work in this new field of objective tests, especially in the
first two years of the study of modern languages. The Department of Romance Languages announces the textbooks and provides a syllabus for each semester of the course. Informal objective tests have been constructed based on the textbooks. These are taken by the students at regular intervals. In addition there are reading requirements and tests based on that. At the present time they have tests based on forty books on their reading list. Students are expected to score at least 80% on these tests in order to go on with the next book on the list. Thus, it can be seen that a very close check is made on every student and all marking is purely objective. Specimen copies of these tests will be found at the end of this paper.

In the field of secondary education, objective tests are being more widely used. The principles of objective tests are being used in the exercises of the more recent textbooks. Workbooks have been published in which all exercises are based on the idea of objective marking. To accompany the French books of Smith and Roberts, there are now available French Progress Tests based on the units of work. The number of schools now using the Cooperative Test Service is increasing annually because this type of test is being more widely accepted.

A criticism of our secondary schools and colleges which is often heard today is that education costs more than it did a generation ago yet the result is an inferior product. Standardized tests in the modern languages lend themselves readily to adaptation for placement purposes.
Ben D. Wood* has shown the misplacement of pupils in the junior high schools of New York, the overlapping of classes in the high schools and the variability of standards. Experiments in other cities would produce the same evidence. There is a great deal of dissatisfaction with the system generally followed of lumping together students whose capacities range from brilliant to bare mediocrity. No one will refute the difficulty of attempting to strike an average pace which will keep the brilliant at least partly occupied and will not be too fast for the mediocre students. Is it not true that many college and secondary school instructors will not bother with students in their classes whose previous functional knowledge is not what it ought to be?

Only standardized tests can be used to conduct the experiments on a scientific basis as those of Cheydleur** at the University of Wisconsin, Tharp*** at the University of Illinois

*New York Experiments with New-Type Modern Language Tests"
The Macmillan Company, 1927.

**The Use of Placement Tests in Modern Languages at the University of Wisconsin" Modern Language Journal, Volume XV, Number 4 January 1931, pages 262-280.

***Sectioning Classes in Romance Languages" same Volume XII, Number 2, November 1927, pages 95-114.
Seibert and Goddard* at Goucher College, and others have proved that these tests for placement and segregation is effective, beneficial and profitable.

New-type tests stimulate pupils to study. Under such a system pupils cannot "spot" questions for an examination because such a test extends over too wide a field. Furthermore pupils who take the "new-type" tests do not enter an examination room with that mental fear which they have for the essay type examination.

The results of the new-type tests can be used to measure pupils so that they may receive more efficient and beneficial instruction. Individual differences are revealed and remedial work assigned for corrective purposes.

The tests will also prove particularly useful in making comparisons between the achievements of different classes in the same or in different schools, at the same time or at different times; between different methods of instruction; between the use of different textbooks; or between the work of different teachers.

From the teacher's point of view these tests are preferred because all scoring is objective, they are more easily and more accurately scored, and the results can be compared with norms established by the authors of the tests.

*"The Use of Achievement Tests in Sectioning Students" same Volume XVIII, Number 5, February 1934, pages 289-298.
IV Conclusion:

It is evident that progress in any field is largely limited by the degree to which accurate measurement is possible and experience indicates that in education as in other fields, the more exact our measurements are, the more surely we shall improve.

Until recent years, most systems of measurement have been very inaccurate. The mortality in our schools is increasing as well as school expenditures. A marked discrepancy exists, and, until this situation is adequately remedied, we, as teachers, are going to find it extremely difficult to justify our existence on account of the pressure brought to bear by the public. Taxpayers are awakening, they want to know how the tax dollar is being spent, and they have a right to know the cost of operating their schools because the school appropriation is perhaps the largest single item of the municipal budget. We must account scientifically for this high mortality.

In the first part of this paper, I have shown that very considerable variation in foreign language standards does exist. I maintain that it is possible to reduce this variation by a uniform marking system and by an application of standards fixed by the actual performance of a large number of candidates who submit to certain examinations.

I have previously proved that most of all ordinary school marks given pupils' work in examinations and oral recitations depend to a great extent upon the person giving them. There is very little agreement in the marks of pupils by different
individuals or by the same individual at different times. With this type of subjective marking and testing such great variability occurs because different teachers have different bases of marking. It is difficult to find two teachers placing the same values on a given subjective test.

I feel that modern language teachers would do well to investigate the "new-type" techniques in examining foreign languages in order to eliminate from our measurement of pupils personal opinion and bias. The amount of time in taking the tests and scoring them is reduced and thus time is saved for other important educational uses. There is also a desirable change in the pupils' attitude as they are not affected by favoritism. Let us work out some uniform marking method and adopt it for our use. Then we shall be able to find out where our standards of achievement are at all levels of instruction. Perhaps thereafter there would be a better understanding between high school and college teachers.

It would not be amiss now to mention that, at the April meeting of the College Entrance Examination Board this year, the new types of examinations adopted by the board have won wide approval because

1 they seemed scientifically sound and
2 they have permitted latitude to students with varied types of preparation. Secondary schools them-selves speak of present methods of scoring and interpreting results which have made it possible to balance factors in final judgment reached in the case of each individual.
Much objection has been raised as to the disadvantage of objective testing:

1. Objective tests measure only memorized material and knowledge of isolated facts. Perhaps there is some truth in this statement but it is possible to construct tests that do more than this.

2. They are difficult to construct. While it is true that objective tests require more thought in preparation, time is saved in scoring and the same test may be used several times.

3. They do not give training in certain abilities in which traditional examinations afford training. In answer to this, I would say that the principal purpose of examinations is not to furnish training in composition but to measure and provide the teacher and pupil with information concerning what the pupils need to be taught as well as what they have learned.

4. They encourage guessing. First in reply to this, I say that guessing is not eliminated in any type of test. Guessing may be penalized in objective tests by subtracting from the total score. Furthermore, is it not common in life to have to choose from among several possibilities?

5. They confuse pupils as to what they know. If the material has been carefully studied the danger of confusion is not apparent. This evokes another real life situation. While I am on this point, is it not true that many teachers are deliberate in their attempt to confuse their pupils in examinations?
The judgment of competent persons is better than the results of tests. I have explained elsewhere that, if supposedly competent people do not agree, there seems little to justify putting faith in their estimates.

Problems in education have increased proportionately with the increase in the number of students in our high schools and colleges. One of the greatest educational problems today is the adjustments of our courses to meet the individual differences of abilities and interests of our students. We must have as much reliable information as can be secured about them, not only for admission to classes but for proper placement.

We must not allow students to waste time on subjects from which they will learn only discouragement. Too often, without regard for the mastery of fundamentals, teachers just "cover" the ground in the texts. They stop for nothing; they never go over a lesson to clarify difficulties. If they would take up no new material until the class as a whole has mastered the fundamental subject matter, they will find at the end of the two years that their pupils will really know something, even though they have not "covered" everything in the book. Students in the modern languages must be carefully and correctly graded and placed in classes within their abilities.

The problem of pupils' marks is very complicated, for what one teacher considers satisfactory work another worthless or a third, superior. At the present time, one of the best instruments of measure available is through the use of objective tests for frequent short testing, and by the use of
standardized tests for final examinations. Testing in this manner eliminates much of the subjective element in the examiner. Pupils are readily made aware of their standing in relation to other members of the class. They can see for themselves that no partiality is being shown to any member of the class. It makes for competition, which may be helpful in raising the level of the work in the class. Furthermore, the teacher can test his own teaching by comparing the norms for the various parts of the tests with the results obtained in his own class. If a teacher finds that his pupils fall below the norm on any particular part of the test, he can readily rectify the error with subsequent classes.

It is possible for any school to institute objective measurement of its work without any additional expense other than the purchase of tests. Enough literature on the subject of "new-type" tests is available for those who wish to construct their own. As for the standardized tests, directions for administering and scoring keys accompany them so that no expert direction or complicated instruments are necessary.

Our methods of teaching have made great advancement, our textbooks have become more scientific, but our examinations remain the same as they were years ago. Let us bring our system of examinations up to date for, the aim of education is identical with the aim of life. Changes in education must keep pace with the changes in our economic, social and mental life.
Pronunciation Test

Matching type

DIRECTIONS: Each of the words in column I has only one vowel sound. Each of the words in column II has a vowel or group of vowels underlined. Each underlined vowel (or group) has the same sound as one of the vowels in column I. In front of each word in column II write the number of the word in column I which as the same vowel sound as the underlined vowel.

<table>
<thead>
<tr>
<th>I</th>
<th>II</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>chien</td>
</tr>
<tr>
<td>2</td>
<td>tante</td>
</tr>
<tr>
<td>3</td>
<td>robe</td>
</tr>
<tr>
<td>4</td>
<td>pas</td>
</tr>
<tr>
<td>5</td>
<td>fille</td>
</tr>
<tr>
<td>6</td>
<td>mère</td>
</tr>
<tr>
<td>7</td>
<td>plume</td>
</tr>
<tr>
<td>8</td>
<td>brun</td>
</tr>
<tr>
<td>9</td>
<td>table</td>
</tr>
<tr>
<td>10</td>
<td>me</td>
</tr>
<tr>
<td>11</td>
<td>vif</td>
</tr>
<tr>
<td>12</td>
<td>soeur</td>
</tr>
<tr>
<td>13</td>
<td>pain</td>
</tr>
<tr>
<td>14</td>
<td>eau</td>
</tr>
<tr>
<td>15</td>
<td>boîte</td>
</tr>
<tr>
<td>16</td>
<td>deux</td>
</tr>
<tr>
<td>17</td>
<td>jour</td>
</tr>
<tr>
<td>18</td>
<td>pont</td>
</tr>
</tbody>
</table>
Grammar Test

Completion Type

DIRECTIONS: Each translation of the following English sentences is correct as far as it goes, but it is incomplete. Decide what is necessary to complete the translation and write it in the space at the right of the French sentence.

1 Does the class understand French?
   La classe ______ le français? (______________)

2 Give her the pencil.
   Donnez ______ le crayon. (______________)

3 It is easy to do that.
   Il est facile ______ faire cela. (______________)

4 They have arrived early.
   Ils ______ arrivés de bonne heure. (___________)

5 We ought to answer them.
   Nous ___ leur répondre. (______________)

6 Are you as tall as he?
   Êtes-vous aussi grand ____ lui? (______________)

7 She will tell you her story.
   Elle vous racontera _____ histoire. (___________)

8 We have not found them.
   Nous ne _____ avons pas trouvés. (____________)

9 That man is our doctor.
   ______ homme est notre médecin. (______________)

10 Are you looking for a fountain pen?
    ________ un stylo? (______________)
11 How old are you?
   Quel âge ______ vous? (__________) 
12 I want soup and roast beef.
   Je désire du potage et _______. (__________) 
13 Have you studied in France?
   Avez-vous étudié _____ France? (__________) 
14 Whose notebook is this?
   ______ est ce cahier? (__________) 
15 What is your name?
   Comment ____________? (__________) 
16 Gold and silver are precious metals.
   L'or et l'argent sont ______ métaux précieux. (_____) 
17 The children are upstairs.
   Les enfants sont _____ en haut. (__________) 
18 Read the fifth paragraph.
   Lisez le _______ paragraphe. (__________) 
19 I get up at half past seven.
   Je ________ à sept heures et demie. (__________) 
20 When he comes send him to my office.
   Quand il ________ envoyez-le à mon bureau. (_____) 
21 Have you enough vegetables?
   Avez-vous _________ légumes? (__________) 
22 The watch my mother gave me runs well.
   La montre que ma mère m'a _____ marche bien. (______) 
23 Is the weather warm?
   ________-il chaud? (__________)
24 If you want some buy them.
   Si vous _____ voulez achetez-les. (_______)
25 We saw her at the theatre.
   Nous l'avons ______ au théâtre. (_______)
26 They used to go to the country every summer.
   Ils ___________ à la campagne tous les étés. (_______)
27 Did I say that?
   ______ j'ai dit cela? (_______)
28 Had he come, I should have gone.
   ______, je serais parti. (_______)
29 What will you buy first?
   Qu'est-ce que vous achèterez _____? (_______)
30 He was among them.
   Il était __________. (_______)
31 How long have you been waiting?
   Depuis quand ________? (_______)
32 There will be about 100 guests.
   Il y aura une centaine _____ invités. (_______)
33 I think it will rain.
   Je pense qu'il ___________. (_______)
34 I know the gentleman but I don't know his name.
   Je ______ le monsieur mais je ne _____ pas son nom.
   (_______) (_______)
35 They have come back from church.
   Elles ________ de l'église. (_______)
36 Do you need this umbrella?
   ________ besoin de ce parapluie? (_______)
37 Who is it?
   Qui ______?

38 It is eleven o'clock.
   Il est _______.

39 We have our books, haven't you yours?
   Nous avons nos livres, n'avez-vous pas______? (_______)

40 Don't lose them.
   ___________.

41 Were you right?
   ___________?

42 I must finish my work.
   Je ______ finir mon travail. (_______)

43 Those are mine.
   ______ sont les miens. (_______)

44 Today is the eighth.
   ______ est aujourd'hui le huit. (_______)

45 He earns five dollars a day.
   Il gagne cinq dollars _____jour. (_______)

46 Aren't you ashamed?
   _______ honte? (_______)

47 She bought a brown leather bag.
   Elle a acheté un sac ___________. (_______)

48 What flowers have you ordered?
   ___________ avez-vous commandées? (_______)

49 Our pens are better than those.
   Nos plumes sont _________ que celles-là. (_______)

50 It is her turn.
   C'est _______.
V Bibliography

Baker, Florence, M; "The Teaching of French"
Houghton, Mifflin Company, 1931

Cole, Robert D; "Modern Foreign Languages and Their Teaching"
Revised and enlarged by James Burton Tharp
D. Appleton-Century Company, 1937

Coleman, Algernon, "The Teaching of Modern Foreign Languages in the United States"
Volume XII-Publications of the American and Canadian Committees on Modern Languages
The Macmillan Company, 1929.

Coleman, Algernon, "Experiments and Studies in Modern Language Teaching"
University of Chicago Press, 1934

Henmon, V.A.C., "Achievement Tests in Modern Foreign Languages"
Volume V--Publications of the American and Canadian Committees on Modern Languages
The Macmillan Company 1929.

Munroe, Walter S. and Souders, Lloyd B., "The Present Status of Written Examinations and Suggestions for their Improvement"

Odell, Charles W., "Traditional Examinations and New-Type Tests"
The Century Company, 1928.
Odell, Charles W., "Objective Measurement of Information"
University of Illinois Bulletin, Volume XXIII, Number 36
May 11, 1926, University of Illinois, Urbana, Illinois.

Rinsland, Henry D., "Constructing Tests and Grading"
Prentice-Hall, Inc. 1927

Ruch, Giles M., and Stoddard, George D., "Tests and Measurements in High School Instruction"
World Book Company 1927.

Ruch, Giles M., "The Objective or New-Type Examination"
Scott, Foresman & Company, 1929.

Several Authors, "Studies in Modern Language Teaching"
Volume XVII--Publications of the American and Canadian Committees on Modern Languages
The Macmillan Company, 1929.

Tharp, James B., "The New Examination vs. the Old in Foreign Languages" School and Society, Volume XXVI Number 674 November 1927 Pages 691-694

Wood, Ben D., "New York Experiments with New-Type Modern Language Tests"
Volume I--Publications of the American and Canadian Committees on Modern Languages
The Macmillan Company, 1927.
Hispania
Bramlage, Julia A., "Factors Affecting Student Progress in Language in the Senior High School"
Volume XIII, Number 4, October 1930 Pages 331-340.
Miner, Glenn B., "The Measurement of Achievement in Spanish"
Volume XIV, Number 6, December 1931, pages 457-482.

Modern Language Journal
Alpern, Hyman, "New-Type Examination in Spanish Literature"
Volume XIV Number 4, January 1930, pages 290-295.
Burchinal, Mary C., "What Should an Examination Disclose as to the Ability of a Student at the End of His High School Course"
Volume I, Number 5, February 1917 pages 163-171
Cheydleur, Frederic D., "The Use of Placement Tests in Modern Languages at the University of Wisconsin"
Volume XV, Number 4, January 1931, pages 262-280.
Cheydleur, Frederic D., "Results and Significance of the New-Type of Modern Language Tests"
Volume XII, Number 7, April 1928, pages 513-531.
Crooks, Esther J., "A Device for Making Tests Function."
Volume XVI, Number 6, March 1932, pages 494-499.
Corbin, Alberta L., "Special Promotions for Superior Students"
Volume XII, Number 2, November 1927, pages 115-117.
Feder-Donald R., and Cochran, Grace, "A Study of Individualized Instruction at the College Level"
Volume XXI, Number 1, October 1936, pages 23-25
Henmon, V.A.C., "Recent Developments in the Study of Modern Foreign Language Problems"
Volume XIX, Number 3, December 1934, pages 187-201.

Lundeberg, Olav K., "Recent Developments in Audition Speech Tests"

Meras, Albert, "French Examinations"

Miller, Minnie M., "A Test on French Life and Culture"
Volume XX, Number 3, December 1935, pages 158-162.

Seibert, Louise C., and Goddard, Eunice, "The Use of Achievement Tests in Sectioning Students"
Volume XVIII, Number 5, February 1934, pages 289-298.

Soldner, Dora M., "A Test in Need"
Volume XXIII, Number 5, February 1939

Stalnaker, John M., and Haden, Ernest, "A New Type of Comprehensive Foreign Language Test"
Volume XIX, Number 2, November 1934, pages 81-92.

Stevenson, Mary Lou, "The Use of Modern Language Placement Tests at the University of Pittsburg"
Volume XVIII, Number 7, April 1934, pages 433-450.

Stoudemire, Sterling A., "Placement Tests in Spanish"
Volume XXI, Number 8, May 1937, pages 593-596.

Sundeen, Myrtle Violet, "Individual Differences in the Modern Language Class"
Volume XIII, Number 2, November 1928, pages 112-117.
Tharp, James B., "Sectioning Classes in Romance Languages"
Volume XII, Number 2, November 1927, pages 95-114.

Tharp, James B., "How Shall We Section Beginning Foreign Language Classes"
Volume XIII, Number 6, March 1929, pages 433-444.

Tharp, James B., "Effect of Oral-Aural Ability on Scholastic Achievement"
Volume XV, Number 1, October 1930, pages 10-26.

Wilson, William E., "Spanish in the Intensive Study Plan at the University of Washington"
Volume XXIII, Number 4, January 1939, pages 257-262.
French Review

Baker, Florence M., "Test Forms for Classroom Material"
   Volume II, Number 1, October 1928, pages 27-45.

Cheydeur, Frederic D., "The Relative Reliability of the Old
   and New-Type Modern Language Examinations"

Giduz, Hugo, "The High Mortality in College Entrance"
   Volume X, Number 6, May 1937, pages 453-460.

Kaulfers, Walter V., and Arata, Isabel M., "Aural Comprehension
   Exercises for Beginners"
   Volume XI, Number 5, March 1938, pages 378-385.

Kurz, Harry, "Mark Hopkins on a French Log"
   Volume IV, Number 2, November 1930, pages 105-125

Price, William R., "Foreign Language Teaching in American
   High Schools"
   Volume II, Number 1, October 1929, pages 17-23.

Tharp, James B., "Examining Knowledge of Foreign Language
   Grammar"
   Volume II, Number 6, May 1929, pages 486-499.

Wadsworth, James R., "They Must Not Fail"
   Volume X, Number 2, December 1936, pages 115-126.
   "They Do Not Fail"
   Volume XI, Number 3, January 1938, pages 228-234.