THE ABILITY OF ENGLISH SPEAKERS TO RESPOND TO THE STRUCTURAL CUES OF WRITTEN LANGUAGE: MEASURING INSTRUMENTS

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

Hal W. Hepler

1966

THESIS

LIBRARY
Michigan State
University

This is to certify that the

thesis entitled

The Ability of English Speakers to Respond to the Structural Cues of Written Language: Measuring Instruments

presented by

Hal W. Hepler

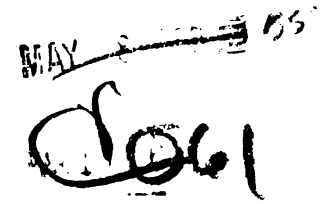
has been accepted towards fulfillment
of the requirements for

Ph.D degree in Communication

Major professor

Date February 14, 1966

O-169


MAY
1961

ABSTRACT

THE ABILITY OF ENGLISH SPEAKERS TO RESPOND TO THE STRUCTURAL CUES OF WRITTEN LANGUAGE: MEASURING INSTRUMENTS

by Hal W. Hepler

Linguists have traditionally defined two kinds of meaning. Lexical meaning is that associated with words and consists of the dictionary meanings of the word. Structural meanings are associated with syntax and consist of such things as knowing which word is the subject of the sentence; when the action takes place; and what the relationship is between the words and word groups of the sentence. We have a number of tests that measure a subject's ability to respond to lexical meaning. We are less fortunate in having tests that measure a subject's ability to respond to structural meaning. This dissertation attempts to measure and index the ability of college age, English speaking subjects to respond to written tests designed to minimize lexical meaning and emphasize structural meaning.

Eight experimental tests were devised and administered to two hundred subjects. Test I asked for judgements of grammaticality and used both real and nonsense words. Tests II and III asked subjects to replace deleted words in sentences from a list of words following each sentence. Test II deleted form-class words; Test III

deleted function words. Test IV, using real English words, asked the subjects to name the parts-of-speech.

Test V asked the subjects to make a translation into "real" English of sentences that were made up of English words that were near homonyms for the words they were to be translated into. For example, Debt's jest hormone nurture was considered to be correctly translated by That's just human nature. Test VI asked the subject to identify a word in one sentence that had the same function as an underlined word in a second sentence. The first 25 words in Test VI were nonsense words and the last 10 were real.

Test VII was like Test IV except that the form-class words, instead of being real, were made-up. The sentence patterns were identical.

Test VIII used the same technique as Test V except that the sentences to be translated were not part of continuous discourse. This was done to eliminate the cues that came from knowing the story being used as a pattern.

Scores were collected for each subject on the College Qualification Test, the Michigan State University Orientation Tests, and on five parts of the Iowa Silent Reading Tests. Each subject reported his grade point average and a number of other demographic variables.

Means and standard deviations were obtained for each variable. Standard errors of measurement and

reliability coefficients were computed for each experimental test. Item analysis was done on the items of Tests I, IV, VI, VII, and VIII using two different techniques. Factor analysis was done on the scores for the eight experimental tests plus twenty other variables. Another factor analysis was done on the experimental tests alone. Correlation matrices were obtained for all variables.

Six of the eight experimental tests had reliabilities high enough for experimental purposes. The twenty-nine variable factor analysis showed six factors accounting for 53% of the variance. The six factors extracted in the twenty-nine variable analysis were tentatively labelled as: (1) a "speed" factor; (2) an "ability to name the parts-of-speech" factor; (3) a "generalized reading comprehension-word knowledge" factor; (4) a "numerical-mathematical" factor; (5) a "word knowledge plus structural meaning knowledge" factor; and (6) a "judgement of grammaticality" factor.

The three factors extracted when the experimental tests were analyzed as a unit were tentatively identified as: (1) an "ability to name grammatical categories" factor; (2) an "ability to respond when certain structural constraints are placed on the response" factor; and (3) a "judgement of grammaticality" factor. The factors of most interest in terms of the present study are 5 and 6 in the twenty-nine variable analysis. These factors are best indexed by experimental tests I, V, VI, and VIII.

THE ABILITY OF ENGLISH SPEAKERS TO RESPOND TO THE STRUCTURAL CUES OF WRITTEN LANGUAGE: MEASURING INSTRUMENTS

by

Hal W. Hepler

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Communication

1966

© Copyright

by

Hal W. Hepler

1966

ACKNOWLEDGEMENTS

This project could not have been undertaken and completed without the guidance and encouragement of a number of people.

My wife and family, in particular, were patient and understanding.

Dr. Erwin Bettinghaus, the chairman of my guidance committee, was unfailingly generous of his time and wisdom. Dr. John N. Winburne gave freely of his wide knowledge of language and linguistics and was instrumental in the preparation of the early materials. Dr. Randall Harrison and Dr. Hideya Kumata gave inspiration and support at all stages of the research.

Dr. Mary Virginia Moore and Dr. Milton B. Dickerson encouraged and supported the research. Mr. Richard Sandow generously aided in the administration of the tests.

TABLE OF CONTENTS

CH. I. INTRODUCTION

This study investigates the responses 200 native English-speaking, college-age subjects made to structural cues in written language. It consists of eight experimentally developed tests plus a number of other measured variables including scores for each subject on a test of reading ability and on the Michigan State University Orientation Tests. In addition, data were gathered on a number of other variables including reading habits, age, sex, high school size, academic major, class, and cumulative grade-point average.

The purpose of the study is, first, to determine if there are differential abilities to respond to tests that purport to measure structural or grammatical meaning, and second, if such differential abilities exist to attempt to measure and index them.

In attempting to provide answers to the above problems a number of psychometric techniques are used. These include factor analysis of item scores, and factor analy-

sis of total test scores. In addition, reliability coefficients are calculated for each experimental test. Validity is determined by correlational techniques and by the use of factor analysis. Item analysis using two different procedures was done on the items of five of the eight experimental tests. The reasons for not doing item analysis on all the tests is explained in Chapter III.

This study was undertaken to attempt to measure objectively the ability of mature, native speakers of English to react to tests which minimize the ability to deal with the lexical element of language. Linguists point out two major dimensions of meaning. One is lexical meaning, consisting principally of synonyms. We have many tests to measure the ability of a subject to provide a suitable synonym for a word. Two such tests are included in the total battery used in this study.

We are less fortunate in having tests that attempt to measure the ability to react to the structural cues of the written language. The existence of these structural cues can be best illustrated with an example. Look at the sentence "The vorpal moggens dapazed molently on the frim." We can make some meaningful statements about this sentence even though five out of the eight words have no referents. For example, most readers of the sentence can say that it is about moggens, that these moggens are vorpal ones, that the moggens did something called dapaze at some

time in the past, and that they did this on something called a frim.

Most of us can respond in this fashion because we have a complete knowledge of the structure of our language. We know, in sentences like the one cited, that subjects appear before verbs. We know that most nouns form plurals by adding -s. We know that modifying adjectives most often appear before the nouns they modify. We know that the will be followed by a nominal.

This study attempts to find out if there are differences in the way native speakers react to these cues. The number of studies of this type are relatively small, and yet research of this kind would seem to be quite important. It is, at least partially, at the syntactic level that we understand and comprehend sentences. The study of the ability of subjects to respond to structural cues should have something to contribute to our understanding of how we read and how we understand what we read.

RELATED LITERATURE, THEORY, AND RESEARCH

In this chapter we will examine two areas of literature, theory, and research: (1) the theory and literature dealing with structural meaning, and (2) other relevant research.

The Theory and Literature of Structural Meaning

Many writers call our attention to the existence of structural meaning. Fries (34:57) reminds us that it is impossible to speak or understand any language without some familiarity with its grammar. This requisite knowledge of grammar is not necessarily conscious, but has to do with one's ability to make proper responses to the various devices that signal the structure of the language and with the ability to produce structural signals that provide the speaker and writer's own structural meanings.

Fries, perhaps the earliest and most influential writer to have concerned himself with structural meaning says:

> The total linguistic meaning of any utterance consists of the lexical meanings of the separate words plus such structural meanings. No utterance is intelligible without both lexical meanings and structural meanings. (34:56)

He continues on the importance of structural meaning:

> Structural meanings are not just vague matters of the context, so-called; they are fundamental and necessary meanings in every utterance and are signalled by specific and definite devices. (34:56)

John Carroll provides a straight-forward description of structural meaning when he points out that structural meanings are those "conveyed by patterns of arrangement and the selection of form classes, as contrasted with lexical meanings, the meanings of the forms themselves." (18:38)

Another writer points out that in language the whole is greater than the sum of its parts. In language as the organization of discourse becomes increasingly complex something new appears which was not present or predictable at an earlier level. Francis illustrates as follows:

> Thus we have seen that when phonemes are organized into allomorphs, they take on meaning, which is not a quality associated with the individual phonemes at all, but is solely a function of the way they are combined. In the same way, when morphemes (or groups of morphemes we call words) are organized into utterances, a new kind of meaning emerges which is not associated with the individual morphemes at all, but is solely a function of the way in which they are combined. This we shall call <u>structural meaning</u>. (30:227)

Lloyd and Warfel (68:98-99) use Lewis Carroll's poem "Jabberwocky" in illustrating how structural meaning operates. They show that while most words in the sentences

of the poem have no referent in the real world the order of the words is the order of English: the structure words are all present, and the phonology signaling structural meanings is what we have come to know, expect, and respond to in English. They continue that if the nonsense words are left out and replaced with blanks, we find a framework into which we can fit our own words. For example, the first line of the poem goes "Twas brillig and the slithy toves did gyre and gimbel in the wabe." If we keep the structure words and the word endings signaling structural meanings we have this: Twas _____, and the _____y _____s did _____ and _____ in the _____. If we want to substitute other words into the slots, we can devise an infinite number of utterances. We could say for example "Twas summer, and the silly boys did splash and swim in the river."

Sumner Ives makes an interesting and clear distinction between lexical and structural meaning. (Mr. Ives' term for structural meaning is grammatical meaning.)

> Whenever we say or write something, the total meaning is compounded from several ingredients. It is something more than the sum of the separate meanings of the individual words. Take the following list of words: fine, does, boy, good, every. Each has a meaning that can be found in a dictionary, but in this order the list is simply a list--nothing more. Now put them into an order dictated by the rules of English structure: "Every good boy does fine." This is an intelligible sentence. In this order the words have grammatical relationships.
>
> Thus the forms of words and the patterns according to which they are put together contribute meaning to the total expression. This aspect of meaning

> which is added to the individual word meanings is called GRAMMATICAL MEANING. (55:27)

Bloch and Trager in discussing structural or syntactic meaning say that just as it is impossible to describe the classes of morphemes (parts-of-speech) except on the basis of form, so also it is impossible to describe adequately the constructions, positions, and form-classes of syntax by talking about their meanings, or by referring to the dictionary or to another language. It is possible to describe the syntactic structures only by reference to their formal, recognizable features--their form and function. However:

> . . . we do want to know something about their meaning. In fact, it is only when we somehow distinguish the meanings of phrases with identical junctural and intonational features that we can recognize the different syntactic types and constructions. (7:74)

Berlo (6:201) in discussing structural meaning says that there is meaning in the form of the sequences of language. He says, "This kind of meaning does not refer to anything, it does not denote anything, but it does aid us in sorting out meanings, in communicating our ideas, and in understanding other people." He further says that knowledge of structural meanings allows us to predict one word-sign from our knowledge of another word-sign. Also, structural meaning allows us to know something about the relationship between two word-signs that we would not be able to know from either word by itself.

The list of writers acknowledging structural meaning

and attempting definitions as significant in linguistics could be expanded for every linguist pays it some attention. Other such writers include: Hughes (53:153ff); Guyer and Bird (46:117ff); Gleason (38:149ff); Hockett (52:137-144); Guth (45:66-71); Roberts (82:171-184 and 83: 58-59); Carroll (17:109). This list by no means exhausts the scholars concerned with the importance of structural signals in language, but they are representative of the whole.

A few have pointed to the research possibilities inherent in the study of structural meaning. Chomsky (21:103) says

> . . . we do find many important correlations, quite naturally, between syntactic structure and meaning; or to put it differently, we find that the grammatical devices are used quite systematically. These correlations could form part of the subject matter for a more general theory of language concerned with syntax and semantics and their points of connection.

Carroll (18:192-193) in commenting on the application of linguistics in measuring psychological traits says:

> Almost all psychological tests are based either directly or indirectly on language responses. It has already been found, by factor analysis and other methods, that some psychological traits are specifically related to the ability of the individual to handle his language. The acquisition of a large vocabulary in one's native language seems to be an index of intelligence. Vocabulary, however, is only one aspect of a language system; it is most nearly like what the linguists call the lexicon on a language. Would it not be interesting to investigate, therefore, the extent to which one might develop measurements of the ways in which individuals handle other aspects of their language systems. The number of investigations which bear on this possibility is so extremely small that the field can be reguarded as unexplored. (My italics)

Lado (62:11) reinforces this view:

> Much remains to be known about how we learn our native language. Detailed studies of individual children report when particular sounds were first uttered and when particular words were first understood and used. In general, these studies are limited by their view of language as isolated words and sounds without relation to the total system of contrasts that is a language. Studies which take into account the relevant structure of language are few. (My italics)

Further insight into the relevance of such investigations appear in the comments of Harris (49:375):

> Finally, there are possible correlations between the descriptive system of a language and investigations in other disciplines. The whole system or features of it may correlate with features of the change and diffusion of language, the formal techniques of the verbal arts, the relation of native speakers to language materials, the processes of language learning, the relation of speech to other human actions, or the relation of linguistics to other sciences.

As can been seen there is a need for structural linguistic studies. It is important to investigate how native speakers respond to structural cues in their language. The findings are valuable in several related areas. It may be possible to learn something about how people read, how they write, how they obtain meaning from the written sentence, how they learn foreign languages, etc.

Related Research

Much of the research reported in this section is peripheral to this study. I have been unable to find research bearing directly upon the problem investigated here. However, many of the factorial studies of verbal behavior give hints that point in the direction taken in

this project.

Rogers (86) attempted to measure verbal fluency. He used a battery of twenty-six tests including such things as grammar (detecting and correcting grammatical errors), verbal meaning (finding synonyms), reading rate, etc. He used a centroid factor analysis and found a first factor consisting of twenty-five of the twenty-six tests. He identifies this first factor as g plus v:ed. These symbols are discussed in Vernon (104) and are identified by him as follows: g is general ability or intelligence, v:ed is the verbal-numerical-educational factor as opposed to the k:m factor that he identifies as practical-mechanical-spatial-physical.

Rogers identifies his second factor as oral facility or ability to give spontaneous oral expression to one's ideas. The third factor is identified as "facility in writing." His fourth factor is of most interest in the present study. Rogers calls it " . . . fluency in dealing with words in which one or more formal restrictions were placed on the response, but little reference was made to its meaning." (86:378) This factor seems to relate to the ability of the subject to perceive language structure. For example, one test having a fairly high loading on the fourth factor was based on the ability of the subject to produce a word when given two of the letters in the word. Another test showing high loading on the fourth factor was one in which the subject had to

form a sentence or meaningful phrase when the first and third word could be any word but when the second and fourth word had to begin with certain letters. For example: _____ y_____ _____ s_____
1 2 3 4
might be correctly responded to by saying or writing "My young son sings."

Fruchter (36), who was also interested in verbal fluency, factor analyzed twenty tests. He identified two and possibly three types of verbal fluency. His second type, of the most interest to us, is "fluency of association for common words where there is some restriction placed upon the response." (36:45) It is possible to see the restrictions and the responses to them as being related to the ability to respond to the structure of one's language.

Johnson and Reynolds (57) factored eight tests of verbal ability plus an intelligence test (the Henmon-Nelson Test of Mental Ability) and a reading test (the comprehension section of the Nelson-Denny Reading Test, Form A). A centroid method of factor analysis was employed and two factors identified. The authors identify the first as those tests depending on a free flow of responses. The second involves those tests wherein the selection of responses is restricted to conform to the requirements of the problem. Another way to view the second factor is as a measure of the subject's ability to respond to the inherent structure of his langauge. Their study suffers

from the relatively small number of variables considered and the single rotation of the factor loading matrix. The results, however, are of interest for the purposes of the present paper. The authors suggest in their summary ". . . there may be two fundamental processes involved in solving all problems: (F) the flow of various acts or responses, and (S) the selection of these responses according to the requirements of the problems. This leads to the hypothesis that individual differences in these processes would be important in determining scores on tests which include the solution of problems." (57:194) (Italics mine)

Davis (26) in preparation for the publication of the Cooperative Reading Comprehension Tests identified nine groups of basic skills for reading comprehension. Multiple-choice questions were devised to measure the skills identified from the literature on reading. 421 college freshmen in Connecticut and Massachusetts took the tests. The tests were factor analyzed using a method described by Kelley. (59) The first factor is clearly word knowledge. The second factor " . . . has its highest positive loadings in the two reading skills that demand ability to infer meanings and to weave together several statements." The factor, most relevant for the present paper, is the fifth, which Davis identifies as " . . . the ability to figure out from the context the meaning of an unfamiliar word or to determine which one of several known meanings of a word is most appropriate in its particular contextual

setting."

Once again it seems they are dealing with the ability of the subject to recognize the structure of the written language and to react appropriately to it. The test with the highest loading on factor five is identified by Davis as: "Ability to select the appropriate meaning for a word or phase in the light of its particular contextual setting. This factor (V) has, incidentally, a very low loading on test 1 which is identified as 'knowledge of word meanings.'"

A.S. Artley (2) points out that there seem to be several factors involved in reading and that these are not correlated highly enough with one another to say that the command of one reading skill guarantees possession of another. He points to two needs in the testing of reading abilities: One, tests involving a number of different types of responses; and two, tests " . . . sufficiently diagnostic to be able to delineate reading needs in the several instructional fields." His point that reading is a complex of factors is relevant to our present purposes.

Harris (48) factored fifteen variables measuring language skills. These included five reading tests, four English tests, one spelling test, and five other tests including length of a composition, number of errors in the composition, ratings of the complexity of the sentence structure, number of punctuation marks used in the composition, and number of different words of more than one syllable used in the composition. The subjects were fifth

grade students in Federal schools operated by the United States Office of Indian Affairs.

I mention the Harris study here only because it is an example of one kind of factorial study that has been done. The results are badly confounded by the relatively low degree of proficiency in English of the bilingual American Indian students.

R.S. Langsam (65) presented twenty-one tests to one hundred freshmen students at Hunter College. The tests included seven parts of the Iowa Silent Reading Tests (used in the present study), two parts of the Minnesota Reading Examination for College Students, the Nelson-Denny Reading Test (two parts), Minnesota Speed Reading Test, Inglis Test of English Vocabulary, and the American Council on Education Psychological Examination for College Freshmen, 1939 edition, as well as the Identical Forms Test of the Tests for Primary Mental Abilities.

Thurstone's (101) centroid method was used in the analysis, and five significant factors identified. Factor one was verbal, factor two perceptual, factor three word, factor four number, and factor five (of most interest for our purposes) "seeing relationships." Factor five involved the following tests: The Paragraph Reading test of the Minnesota Reading Examination, the Same-Opposite test of the ACE as well as the sentence meaning test and the paragraph meaning test of the Nelson-Denny Reading Test. Langsam says " . . . the common character-

istics (of the tests loading on factor five) seems to be that of seeing relationships among the elements of the problem . . ." This would indicate some relevance to the problems of structural **meaning** as investigated in the present paper.

Stolurow and Newman (96) using an intercorrelation matrix prepared by Gray and Leary (40) reduced the matrix to a 23 variable matrix and then factor analyzed the matrix using the principal axis method and the quartimax rotation. Factor one was identified as an easy vs difficult word factor. Factor two (of most interest in the present study) was identified as an easy vs difficult sentence.

The authors say: " . . . the 23 elements of expression that are related to reading skill can be roughly grouped under two main factors which we have labelled as an easy vs difficult word factor (semantic difficulty) and a difficult vs easy sentence factor (syntactical difficulty). This of course, comes as no great surprise. It is felt, however, that these results have provided information about the relative importance of structural elements of expression which several investigators have used in the development of 'readability' yardsticks . . ."

The authors continue that there is some evidence there may be differential abilities among readers to react to the factors described and that this finding has implications for the preparation of reading comprehension tests.

as well as readability yardsticks.

It should be made clear that what Stolurow and Newman were analyzing were not (as in the present study) responses to language but to the various elements or particles or structures in the written language. Their study points the direction taken by the present one.

Clark (22) shows that words whose meanings are known to the receiver are "combined in known sentence structures to reveal relationships not known previously. To understand an utterance then, a decoder must determine (1) the structural relationships among the words and (2) the distinct sense of each word as the encoder intended it."

Clark tested the influence of syntax and reference on the ability of subjects to predict unknown words. Her major dependent variable was the ability of the subjects to predict words that had been deleted from messages. This technique is derived from the "cloze" procedure work of Taylor. (98)

The major independent variables were syntactic and referential information. Syntactic information was provided by giving the subject the part of speech of the deleted word; referential by allowing the subject to view the painting serving as the stimulus for the composition of the original message. Her findings, using as analysis of variance design, showed that furnishing syntactic information did increase the number of correct responses.

Providing referential information also increased the number of correct responses.

Her study has relevance to this one in that Clark anticipates some of the present findings. For instance, she finds that native speakers have a very strong, built-in knowledge of their language's structure as evidenced by her report that even when the grammatical class of the unknown word is not supplied, the subjects respond correctly 84% of the time.

Clark also found language habits that allow decoders to interpret structural relationships and semantic clues are relatively common among all native speakers, hearers, readers, and writers. The relatively is the subject of this dissertation.

Miller and Selfridge (73), in a now classic study, report on the influence of contextual restraints on the ability of subjects to recall strings of words. They show that when approximations to English are of the third or fourth order the ability of subjects to recall is about the same as for textual material. An example of a 0 order approximation to English is "byway consequence handsomely financier bent flux cavalry swiftness weather-beaten extent." (p.184) An example of a third order approximation is "tall and thin boy is a biped is the beat." (p.185) The authors found "meaningful material is easy to learn, not because it is meaningful per se, but because it preserves the short range associations

that are familiar to the Ss." In the context of this dissertation it is suggested that these "short range associations" are comparable to the structural meanings previously discussed.

Carroll (16) did a large scale factor analysis that comes, perhaps, as close to the present study as anything reported. Carroll factor analyzed 42 tests using 119 college adults as subjects. The factor of most interest for our present purposes he labels as C or C'. The 15 tests with significant loadings on this factor include: Word-choice, Vocabulary, Phrase Completion, Grammar, Memory for Homophones, Rhyming, Spelling, Morpheme Recognition, Disarranged Morphemes, Theme-Rating, Disarranged Words II, Paragraph Memory, Distorted English, Suffixes, and a Speech Attitude Scale.

A number of the above tests would seem to measure to some degree the subject's ability to react to written language when certain constraints (that is, structural limitations) are imposed on the responses. One example is the Distorted English Test. Here, the subject was given a sentence in Hungarian that has been translated literally, word for word, into English. The subject's task was to render the sentence into idiomatic English. It would seem this test is one way to determine the ability of subjects to respond meaningfully to the structural constraints of English.

In his discussion of the C factor Carroll has this to

say:

> Close examination of the data available leads the writer to conclude tentatively that this factor represents the individual differences in some aspect of the ability to learn various conventional linguistic responses and to retain them over long periods of time. The factor represents differences in the stock of linguistic responses possessed by the individual--the wealth of the individual's experience and training in the English language. By conventional linguistic response may be understood any fact of speech behavior which is essentially arbitrary but which occurs with a certain frequency in definite situations. A response (e.g., the response underlying a phoneme) may not even have any intrinsic semantic value, though most linguistic responses do have such a value. The concept of conventional linguistic response described here is exemplified by words and meanings of words; phonological, morphological, and syntactical features of the language; certain expressive gestures; and patterns of idiomatic expression. (The writer assumes that formal characteristics of a language correspond in some way to responses in a psychological sense.) (16:293)

Two studies are particularly relevant to the method used in Test I of this dissertation. The studies are those of Maclay and Sleator (71) and Hill (51). In both papers the authors are concerned with the question of "grammaticality or grammaticalness."

Maclay and Sleator presented 36 sentences to 57 undergraduate students at the University of Illinois divided into six types: Type A-not grammatical, not meaningful, not ordinary; Type B-grammatical, not meaningful, not ordinary; Type C-not grammatical, meaningful, not ordinary; Type D-not grammatical, meaningful, ordinary; Type E-grammatical, meaningful, ordinary; Type F-grammatical, meaningful, ordinary.

Examples of each type follows: A: A keeps changed very when; B: Appointments can mow winters generously; C: Yesterday I the child a dog gave; D: In order to get there before they close; E: It's better to walk than running; F: He was ready to go. Types E and F are both, according to the criteria, grammatical, meaningful, and ordinary. Type E, however, violates some of the "traditional" as opposed to "linguistic" notions of grammaticality.

The dependent variable, was, of course, the judgments of the subjects on the grammaticality, meaningfulness, and ordinariness of the sentences presented. The findings indicate that while the subjects agreed fairly well with the a priori classifications of the authors there was a significant amount of individual variation. This result agrees with the findings of Test I in the present study.

Maclay and Sleator point out: "The fact that three out of 21 subjects judged the sequence Label break to calmed about and to be grammatical indicates that very little can be assumed in advance about responses to language, and even the most obvious predictions need to be checked empirically." (71:281-282)

Hill, in a similar study, chose eight sentences drawn from examples used by Chomsky. (21) The test sentences were of the following form:

"1. Colorless green ideas sleep furiously

2. Furiously sleep ideas green colorless

3. Have you a book on modern music?

. . .

8. I saw a fragile of." etc.

Hill's ten subjects included one secretary, one undergraduate business major, two linguists, and six college professors of English literature and composition. They were asked to reject ungrammatical sentences and accept grammatical ones. He found quite a wide variance in the judgments of his subjects. This also is in agreement with the findings of Test I in this dissertation. Hill's brief study can be criticized on the grounds that his list of test sentences was quite short, his sample of subjects small, and his selection of subjects included 80% who were by no means naive. The study does show the value of such a procedure and that there are individual differences in the way that subjects respond to the structural cues of their language.

The studies of Berko (90) indicate that children develop rules for dealing with words quite early in life and that there is a progressive increase in the child's ability to deal with the problems she set for her subjects.

In her experiment children ranging from four to seven, including both first grade and preschool children, were presented with nonsense words and asked to supply the English plurals, verb tenses, possessives, etc. of these words. The technique used was to present the child with

a picture of a fanciful animal and say to the child: "This is a wug." Then, the first picture--showing only one of the animals is removed, and a second picture showing two of the animals is presented. The experimenter then says: "Now there is another wug. There are two ___?___" The experimenter uses intonation to show another word follows at the indicated blank and records the child's response. The expected response is wug with /z/ affixed to indicate the plural.

The study found that children were generally able to perform this task and that their ability to solve the problems correctly increased with age and with exposure to formal education. Another finding was that certain inflectional features of the language were learned more slowly than others. For instance, the past tense of bing was given correctly as binged by 60% of the pre-school children and by 85% of the first-graders, while the past tense of ring was given correctly as rang by only 25% of the first-graders and by 0% of the pre-schoolers.

It is interesting to speculate on whether this difference in ability to handle the structural elements of language persists on into adulthood. If the differences in ability to deal with structure do persist it seems reasonable to expect that they will be relatively much smaller than in childhood, but it should be possible to make some steps toward indexing them. This is, at least

partially, the purpose of this dissertation.

Studies by Brown (11) and Brown and Bellugi (13) offer further insight and research into the sequences involved in the child's learning of the structure of his language and reinforce the contention that it is possible that there are individual differences in the ability of speakers to respond to structural cues in language.

CH. III. RATIONALE, PROCEDURE, AND RESEARCH DESIGN

Overview

This dissertation consists of a factor analytic study designed to answer the question: Can we index and measure the ability of native speakers of English to respond to the structural cues inherent in the written language? To aid in answering this question many tools of the psychometrist are used. These include item analysis, correlational procedures, reliability coefficients, factor-analysis of major variables, and factor analysis of item scores.

In this chapter the rationale for the study, the procedures used, and the research design will be discussed. Under rationale will appear what was done and why it was done; under procedure, the subjects, the experimental tests, and the other measured variables; under research design, the various statistical techniques employed.

Rationale

If there is structural meaning as opposed to lexical meaning, it is reasonable to suppose that the subject's ability to respond to it can be measured and indexed. In

an attempt to establish measurement eight experimental tests were devised and administered to discover if any or all of the experimental tests could fairly claim to measure the ability to respond to structure. The tests took various forms: one tested the ability of students to make judgments about the grammaticality of sentences when most of the lexical items had been removed and replaced with possible but non-existent words; two asked subjects to replace missing words in English sentences; a fourth asked the subjects to name the traditional "parts-of-speech" in English sentences to measure the extent to which the ability to respond to structure is a measure of the ability to name the categories of the words in the sentence; a fifth asked the subjects to do a kind of translation from sentences involving near-homonyms of words into "real" English; a sixth asked the students to identify a word in one sentence that had the same function as another word in a paired sentence--most sentences involved possible but non-existent words to minimize the lexical element; that is, to decrease the subject's reliance upon "dictionary" meaning; the seventh asked the subjects to identify the "part-of-speech" of underlined words in sentences in which the words were possible but non-existent words; the eighth was much like the fifth in that the same translation procedure was used. In test VIII the contextual cues that accompanied Test V were removed to

make the test as much as possible a valid measure of the ability to respond to the structure of the sentence rather than a measure of how well the subjects recalled the story.

In addition to the scores on the eight experimental tests, scores were also obtained on seven tests in the Iowa Silent Reading Tests (Form Dm), for the three parts of the College Qualification Tests, and for the Michigan State University Arithmetic, Mathematics, Reading, and English Tests. Data were obtained for each subject on his reading habits in number of books, magazines, and newspapers read regularly, and on his grade point average, academic major, age, high-school size, etc.

Procedure

In this section will be discussed the subjects, the testing procedures, the eight experimental tests, the other tests, and the various kinds of demographic data collected.

The Subjects and the Testing Procedures

The subjects were 200 students enrolled in business writing at Michigan State University during the spring of 1965. Each subject did not participate in all tests, and data was incomplete for some subjects.

There were 31 females and 169 males whose mean age was 21 years and 7 months, ranging from 36 years to 19 years and 6 months.

The subjects came from a wide variety of academic majors as summarized in Table 1.

The various tests were not required of the students, but all those involved cooperated willingly. The testing required approximately three and one-half fifty minute class periods spread over three weeks.

The attitude of the subjects to the various tasks was good throughout although there were some amused comments about the value of dealing with the nonsense words, and some questions about the reasons for the research. Explanations about why the research was being done, and the relevance of the various tests were made throughout the testing. Each part of the testing procedure was kept as informal as consistent with good testing. Explanations were made when requested, and a real attempt was made to be straightforward. There was no "mystery" about the testing process or the aims of the research.

The testing in four of the seven sections was partially handled by Mr. Richard Sandow, the instructor of those sections. The Iowa Silent Reading Tests, requiring strict timing, were done in all sections by the experimenter.

The subjects were all native speakers of English, including 2 sophomores, 118 juniors, 72 seniors, and 2 special students. Both of the special students had earned Bachelor's degrees at other institutions and were attending Michigan State to earn a Bachelor's degree in another field.

The Eight Experimental Tests

To attempt to measure the ability of subjects to respond to the structural cues of written language eight tests were experimentally developed and make the major interest of this dissertation. Five tests were developed first and three more were added as the preliminary results of the first five became known. Each of these tests will be discussed in detail below. Each was a paper and pencil test requiring that the subjects mark their responses directly on the test booklet. All tests with the exception of the first one were scored simply as number right. Experimental Test I was scored as rights minus wrongs.

Experimental Test I. This test consists of twenty-five English statements using real words and made-up words. The subject read each sentence and made a judgment as to whether the sentence could be an English sentence. The student was asked to judge on the basis of whether or not the sentence has the structural pattern of English sentenses. The subject is given two alternatives: English? __ or Not English? __.

The sentences used were drawn from an essay by A.J. Liebling, "Ahab and Nemesis," appearing in The Sweet Science, and reprinted in a collection of essays edited by Leslie Fiedler. (29:320-332)

To provide an appropriate context for this test the

introductory sentence of the essay is provided at the head of the test. (See Appendix III) The sentences following are drawn in order from the succeeding sentences. The changes made were to shorten long sentences to facilitate the administration of the test.

Of the twenty-five sentences, sixteen were considered "English" in structure, and nine "non-English." This was established by having the sentences examined by several professional colleagues serving as sophisticated informants. The inter-judge reliability was very high.

Sentence two consisting of English words was: "In the roily down, Kansas and Leonard." This had a prepositional phrase followed by two proper nouns joined by a conjunction with no verb. The two proper nouns are both mentioned in the introductory paragraph that sets the context.

Sentence three: "Broun was purloinedly refrected" contains some unfamiliar lexical items, but Broun, the subject is identified from the introductory paragraph; was is an English verb; purloinedly is an adverb; and refrected, while a nonsense word, contains the preterite verb structural morpheme - ed. This sentence has the structure of English.

Inasmuch as the subject has only two foils (English? ___or Not English? ___) in response to each question the total score was calculated on a right minus wrong basis to minimize the effect of guessing. Such scoring does not

change the rank ordering of any of the subjects. It lowers the mean score and increases the amount of variance of the total test scores.

This test has its origins in the suggestions of Chomsky (21) and Roberts (81), as well as in the researches of Maclay and Sleator (71) and Hill (51). The hypothesis is that those students having the best command of the structure of their language will have the highest test scores. The lexical element is minimized by providing "nonsense" words in place of many form-class words. The contextual and structural features of continuous discourse are retained by providing an introductory sentence and by presenting the test sentences in the order used by the author.

Experimental Test II. In this test the subject is presented with sentences in which the form-class words, (nouns, verbs, adjectives, and adverbs) have been replaced with blanks. The subject is to select a word that will fit appropriately into each blank from a list of words following each sentence.

Sentence one was: "In the __1__, a __2__ __3__ed __4__ly." Eight words were offered as possible choices: "1. of; 2. be; 3. summer; 4. for; 5. boy; 6. some; 7. slow; 8. walk." The expected replacements were summer for slot 1, boy for slot 2, walk for slot 3, and slow for slot 4.

Sentence three was: "The ____est ____ ____s ___ish
1 2 3 4

with the foils: "1. with; 2. act; 3. fool; 4. girl; 5. were; 6. tall; 7. very; 8. should." The expected answers were tall for slot 1, girl for slot 2, act for slot 3, and fool for slot 4.

In sentence one the structural cues include word order, the preposition in followed by the determiner the signaling a noun follows, the morpheme suffixes -ed and -ly, and the determiner a indicating a following nominal.

The test consisted of ten sentences with forty responses. Two sentences required three slots to be filled, six sentences required four, and two sentences required five. This test, as were all the other experimental tests, was untimed. Having all the experimental tests untimed eliminates the spuriously high reliabilities associated with strictly timed tests. (see Crunbach, p.141-142)

The hypothesis being tested here is that those subjects who are most sensitive to the structure of their language will be best able to replace omitted words and get the highest scores.

Experimental Test III. This test is much like the preceding one, except that instead of blanks the deleted words were replaced with various geometrical symbols. Also, the deleted words, instead of being form-class words, were function words. (The terms form-class word and function word are Fries'. See The Structure of English,

1952) Function words are prepositions, articles, subordinators, and conjunctions that help provide sentence structure. This test examines the subject's ability to replace structure words, as contrasted to Test II which tested for ability to replace form-class words.

Sentence B was like this: "I did not think ▭ he could bring it off, △ I wanted to be there ▽ he tried."

▭ stands for ____.

△ stands for ____.

▽ stands for ____.

Answers: 1. when; 2. by; 3. to; 4. but; 5. up; 6. around; 7. down; 8. that.

In sentence B the expected responses were ▭ equates for that, △ but, and ▽ when.

Six sentences were used and from these sentences twenty structure words were deleted. The sentences were drawn from the Liebling essay previously mentioned.

The hypothesis is that those subjects most familiar with the structure of their language will be best able to replace the deleted words. A second hypothesis was that the ability to replace structure words differs in some significant way from the ability to replace form-class words.

Experimental Test IV. This test measures the ability of subjects to identify the "part-of-speech" of underlined words in sentences. Two sentences were used, both drawn from the Liebling essay. In the first sentence,

consisting of nineteen words, the subjects were asked to identify all of the words. In the second sentence, consisting of fifteen words, the subjects were asked to identify six underlined words for a total of twenty-five.

The second sentence was: "Dempsey may have been (20) a great (21) champion, (22) but (23) he had (24) less to beat than Marciano. (25)" The student is asked to identify each of the underlined words as noun, pronoun, verb, adverb, adjective, conjunction, or preposition.

The twenty-five words to be identified consisted of seven nouns, one pronoun, six verbs, six adjectives, two conjunctions, and three prepositions.

The test was included in the battery to investigate the extent to which success on various of the other tests is related to the ability to attach labels to the words in the sentence. The test attempts to measure the ability of subjects to sort words into categories using a system that is often taught in the elementary and high schools and to investigate the relationship between this test and another (Test VII) in which the second test asks the subjects to identify the part-of-speech of nonsense words used in the same pattern as in Test IV. Test VII is discussed in detail below.

Experimental Test V. This test presents the subject with sentences in which the words are real English words that sound somewhat like other English words. The sub-

ject is asked to translate the "funny" sentences back into real English sentences. For example, the student is given something like this: "Marry hatter ladle limb, itch fleas worse widest snore." This is correctly translated as "Mary had a little lamb, its fleece was white as snow."

The scoring of this test was simplified as much as possible by providing the student with the appropriate number of blank spaces for his translation and by indicating to him those words that were to be translated as two or more words. For example, in the phrase "pimple orphan colder . . ." the correct translation is "people often called her . . ." Pimple translates into people, Orphan translates into often, and colder translates into called her. The fact that colder requires two words in the translation is indicated to the subject by providing him with two blanks below the word.

The untimed test consists of ten sentences with 121 words to be translated.

To provide context for the test an introductory paragraph is offered at the beginning of the test. The test tells in the usual order the first part of "Little Red Riding Hood."

The hypothesis is that one able to make correct translations from the original must know English structure. Even if he is able to list words that are near homonyms for the words in the test sentences, he will be unable to make a translation unless he is able to choose those that

"make sense," a part of the ability to respond to structural cues.

Experimental Test VI. In this test the subject is presented with two sentences of approximately equal length. The second sentence of the pair has each word numbered for identification. The subject is asked to indicate the word in the second sentence which has the same function as an indicated word in the first sentence.

The first pair of sentences and the questions about them are as follows:

I. A. In the frammis, a morgrant arablint daskaped.

B. The borpal toggen profrumes the glaspart.
1 2 3 4 5 6

1. What word in sentence B has the same function as morgrant in sentence A? ____

2. What word in sentence B has the same function as daskaped in sentence A? ____

The test consists of ten pairs of sentences with thirty-five questions. The first seven sentences and the first twenty-five questions use "nonsense" words. The last three sentences and the last ten questions use English words allowing us to compare the results on the last ten items with the results from the first twenty-five.

Two of the pairs used question patterns; one a complex-sentence pattern; the remainder were simple-sentence patterns.

The hypothesis is that the ability to respond

accurately on such test materials is a measure of the ability to respond to structural cues. The use of the comparison technique (what is the word in sentence B that has the same function as word X in sentence A?) allows the subject to respond without having to be able to name the various parts-of-speech.

Experimental Test VII. In this test the subject is again asked to name the part-of-speech of underlined words. The difference between this test and Experimental Test IV is that in this test the words are "nonsense" items. To retain the structural features of English the structure words were retained in their regular form. The first two sentences were copies of sentences A and B in Experimental Test IV with the same pattern, the same structure words, the same affixes, and the same order as the two sentences in Test IV. Test VII adds two additional sentences making a total of twenty-five.

The subjects were tested only on their ability to name correctly the "nonsense" words. The only allowable answers were noun, verb, adjective, and adverb. The subjects were not aware of this limitation, however, and all the foils provided in Test IV were provided in Test VII.

The hypothesis is that one familiar with the structure of his language will find it as easy to name the form-class words when they are "nonsense" as he does when they are "real." If we categorize words on the basis of their

position, on the basis of their affixes, and on the basis of the function words that mark them, we should be able to respond as well in naming parts-of-speech to the "nonsense" items as to the "real" lexical items of Test IV.

Experimental Test VIII. This test uses the same translation technique employed in Experimental Test V. Instead of having a story told in a continuous context, however, Test VIII offers seven unrelated sentences to be translated.

This test was added after the preliminary scoring was done on Test V. It seemed that Test V was too easy for subjects quickly realized that the tale was Little Red Riding Hood, and that part of what was being tested was the subject's ability to recall the story. To minimize the effect of memory, the same procedure was used and the contextual cues were removed. There was no introductory sentence, and the seven test sentences have no relationship to one another.

The first sentence to be translated was: "Debt's jest hormone nurture." The second sentence was: "Oil ketchup wetter letter." The first was to be translated as: "That's just human nature." The second required: "I'll catch up with her later."

Planning the test in this way makes each sentence independent of the others and forces the subject to attend to the cues inherent in each sentence rather than

discourse structure.

This test required sixty responses in contrast to the 121 of Test V. As in Test V a blank is provided for each word required in the translation and credit was given for alternate interpretations. For instance, in sentence two credit was given for it, her, or you in the phrase "with her later."

Experimental Tests I-VIII Total. A total score was derived for the responses to each of the eight Experimental Tests. There were a total of 351 responses.

The Iowa Silent Reading Tests

Each subject was asked to complete seven parts of the Iowa Silent Reading Tests: Form Dm. (42) This test was chosen because it offers a number of separate tests designed to measure various aspects of the ability to read. It was felt desirable to measure as broad a range of reading skills as possible. If, as has been suggested previously (2), reading is a complex of behavior, it would seem wise to use a reading test covering as wide a range of skills as possible.

These tests, as well as the other tests to be discussed (such as the College Qualification Tests and the various Michigan State University Orientation Tests) were included to validate the Experimental Tests; that is, the Experimental Tests must be checked against tests of known validity.

The scores obtained from the Iowa Silent Reading Tests include: a reading rate for Test 1-Rate-Comprehension-Part A; a reading rate for Test 1-Rate-Comprehension-Part B; a reading rate for Test 1-Rate-Comprehension-Part A plus Part B; and a comprehension score for parts A and B of Test 1. In addition, a score was obtained for Test 2, Directed Reading; Test 4, Word Meaning; Test 5, Sentence Meaning; and Test 6, Paragraph Comprehension. Each of these tests is discussed in detail below.

Test 1. Rate-Comprehension--Part A. A test on a 399 word essay (24 sentences) on iron, which the subject is directed to read carefully. After one minute the subject is stopped and records the number of the sentence he is reading. This provides a rate-of-reading score. The highest possible score is 24, indicating the subject had read all of the essay in one minute. The lowest score is 1, indicating that the subject was still reading within sentence 1 at the end of one minute. The scores are not converted into words-per-minute.

Next the subject receives two minutes to complete his reading of the essay, following which the subject has two minutes to answer ten multiple choice questions based on the essay. He is not allowed to review the essay while answering the questions. All answers are recorded on machine-scored answer sheets supplied by the test publisher and machine scored by the Office of Evaluation Services of Michigan State University.

Test 1. Rate-Comprehension--Part B. This test is much like the previous one, consisting of a 633 word essay on unemployment insurance and containing 30 sentences. The procedure for determining the rate of reading is as in Part A. At the end of one minute the subject is stopped and records the number of the sentence he is then reading. He then receives two more minutes to read as much of the rest of the essay as he can. At the end of two minutes the subject is given three minutes to answer twenty-five true-false--not discussed questions about the essay.

The sentences in Part B average somewhat longer than those in the essay of Part A. Average sentence length in Part A is 16.66 words, and average sentence length in Part B is 21.13 words.

Test 1. Rate-Comprehension--Parts A and B. Test 1 of the Iowa Silent Reading Tests then yields four scores: reading rate on Part A, reading rate on Part B, reading rate on Part A plus reading rate on Part B, and the comprehension score on Part A plus the comprehension score on Part B. There are ten questions over Part A and twenty-five questions over Part B so the total comprehension score can range from zero to thirty-five.

This test measures the subject's ability to understand what he reads and provides an index of the relationship between ability to deal with structure and ability to comprehend written material, or it provides one method of

validating the experimental tests.

Test 2. Directed Reading. This test consists of an essay on cork plus twenty questions about the facts in the essay. In answering the questions the subject is directed to show the sentence number containing the answer to the question. There are 24 numbered sentences in the essay.

The Manual of Directions for the Iowa Silent Reading Tests (43) says: "This part of the test is designed to measure the pupil's ability to comprehend general and specific situations expressed in the content without unduly stressing memory. While this test is designed to measure the ability to comprehend questions of a rather detailed type, it makes a special effort to avoid pure identification or matching of words." (43:2) It seems reasonable to suppose that the ability to respond accurately to structural cues has some relationship to the ability to do well on this test. Given a question about the text that must be answered by providing the number of the sentence in the text that gives the answer should put a premium on the ability of the reader toreact quickly and accurately to the sentences provided as foils. The student-subject good at responding quickly and accurately to the structural cues of the written language should be able to say quickly and with few mistakes what the sentence is about.

This test, in common with all of the Iowa Silent

Reading Tests, is strictly timed; the subject is given three minutes to complete the work.

Test 4. Word Meaning. This test, measuring the subject's vocabulary, consists of 70 words, each of which has five discrete words as foils. The task is to choose the foil that provides a synonym for the test word. Some of the test words are offered in a brief context:

> 1. To blockade an enemy is to-- 1. destroy ships 2. bomb troops 3. prevent entrance of supplies 4. lay mines in harbors 5. withdraw credit

Other questions in this test omit the contextual cues. For example:

> 19. Urban refers to-- 1. country 2. something changing 3. census reports 4. city 5. agriculture

This test is divided into four parts: part A consists of twenty words drawn from the literature of the social sciences; part B of fifteen from science; part C of fifteen in mathematics; and part D twenty in English studies. In this paper the seventy word total is treated as a whole; no effort was made to break the total test score down on the basis of the four sections.

Each subject is required to do some part of each of the sections. Seven minutes are alloted for the whole test, but two minutes are given to Part A, one and a half minutes to Part B, one and a half minutes to Part C, and two minutes to Part D. Thus the subject's vocabulary for the areas covered in each of the four parts is sampled.

This test is included because it is a typical reading test measure of the subject's ability to respond to lexical meaning. It allows us to investigate the relationship between tests designed to measure the subject's ability to respond to structural cues and his ability on tests that measure lexical meanings.

Test 5. Sentence Meaning. This test consists of fifty sentences. The subject is to read each sentence and respond Yes or No. The first sentence is: "Are the opinions of experts valuable in certain situations?" The expected response is, Yes. The authors of the tests say that "the sentences comprising this test are stated in such a way that in each case the meaning of the sentence as a whole must be comprehended." (43:2) This test is strictly timed (four minutes are allowed for its completion), and the subjects are put under some pressure by being reminded in the instructions not to guess.

This test is used because it is hypothesized that the ability to respond quickly and accurately to such questions should have a positive relationship with tests designed to measure the subject's ability to deal with structural meanings. The ability to respond to the sentence should be a function of the ability to deal with the structural cues of the written language.

Test 6. Paragraph Comprehension. This test consists of twelve paragraphs of from 70 to 90 words each. The student is asked to read each paragraph and answer three

questions about it. The questions are multiple-choice in form, and each question has three foils. The questions for each of the twelve paragraphs follow the same form. The first question asks the subject to choose the best title for the paragraph. The second and third questions ask the student to demonstrate his understanding of the subject matter of the paragraph. An example of the type of questions asked follows:[1]

A. Choose the best title for the paragraph. 1. Fishing, Ancient Industry 2. Early Fishing Trips 3. Deep-Sea Fishing
B. Why did fishing cause the exploration of distant lands? 1. fishing led man to sail the seas 2. seaports became large cities 3. man has fished for many centuries
C. Many of our large modern commercial cities had their beginnings as --- 1. centers of agricultural activity 2. fishing villages 3. castles of the noblemen

The subject is given nine minutes to read the twelve paragraphs and answer the thirty-six questions.

This test was included by the authors of the Iowa Silent Reading Tests because they felt that it measured two specific aspects of paragraph comprehension. These were: the ability to select the contral topic of the paragraph, and the ability to identify details to understanding the meaning of the paragraph.

Michigan State University Orientation Tests

Orientation test scores were obtained for most subjects involved in this study. With some exceptions all

students entering Michigan State University are required to take a battery of tests. The student's scores on these tests are converted into percentiles and reported to his advisor and his academic dean as well as to others who have an interest in the student. Raw scores for each student are filed in the Office of Evaluation Services and are available for research. These raw scores were used in this study.

Some scores are, of course, not available. The total number of observations available for each test is reported later in this dissertation.

There five orientation tests. These are: The Michigan State University English Placement Test, The Michigan State University Reading Test, The Michigan State University Arithmetic Placement Test, The Michigan State University Mathematics Test, and The College Qualification Tests. Each of these will be briefly discussed below.

The MSU Reading Test. This test is a 42 item test of reading comprehension. The student's score is based on his ability to answer questions based on reading passages that are representative of several academic areas. The score is supposed to provide a measure of the factors involved in critical thought. (99)

The scores on this test are included as part of this study because they provide one more measure against which the validity of the experimental tests can be checked.

The MSU English Placement Test. A 35 item objective test over various aspects of English usage, it includes spelling, capitalization, grammar, punctuation, sentence structure, and organization. The test is designed primarily to identify students who may need remedial help in English although it is also used in identifying those students who might be candidates for honors sections of courses. Again, the test provides an external criterion against which to check the validity of the eight experimental tests.

The MSU Arithmetic Placement Test. This 40 item test of ability in elementary arithmetic is used to detect those students deficient in basic arithmetic who need remedial work.

The MSU Mathematics Test. This 30 item test dealing with high school algebra is used as a predictor in determining whether a student will be successful in technical courses and in placement.

Entering freshmen have an option as to which of the two tests (Mathematics or Arithmetic) they will take. In the group of students used in this study most chose the MSU Mathematics Test. One hundred and twenty took the Mathematics Test and sixty-two the Arithmetic Test. Eighteen took neither.

College Qualification Tests---Verbal Section. This test contains 75 vocabulary items, fifty of which require

identification of synonyms, and 25 identification of antonyms. The test questions are of the following form:

> Choose the word which means the SAME or most nearly the same as the word at the left.
> 1. MUMBLE 1. Huddle 2. Choke 3. Mutter 4. Drop
>
> Choose the word which means the OPPOSITE or most nearly the opposite of the word on the left.
> 2. SILENCE 1. Terror 2. Noise 3. Beauty 4. Warmth

The test is timed, and the student is given 15 minutes to complete it.

The authors (5) claim that the test has quite high predictive validity. They cite validity coefficients of from .19 to .57 with first semester grade point averages in public four-year institutions. This section of the College Qualification Tests also correlates highly with the Nelson-Denny Reading Test and with the Comprehension section of The Cooperative Reading Test. The figures, using the corrected odd-even score coefficient are .78 with the Nelson-Denny and .85 with the Cooperative.

College Qualification Tests---Information Section. This section of the CQT contains 75 items drawn from a broad range of subject matter areas. Half the questions are from science, and half from the social sciences. The test is timed and 30 minutes are allowed for completion. As in all sections of the CQT the emphasis is on power rather than on speed. A typical question takes the following form:

> A. Florida is a 1. plateau 2. delta 3. penin-

sula 4. savannah

This section of the CQT has validity coefficients ranging from .27 to .63 with first semester grade point averages in public four-year institutions.

College Qualification Tests--Numerical Section. This section of the CQT consists of 50 items drawn from arithmetic, algebra, and geometry. The test, according to the authors, measures conceptual skills in simple mathematics rather than computational or clerical skill and speed. Two typical questions follow:

1. 7.064 - .646 = 1. 6.328 2. 6.418 3. 6.994 4. 7.004
2. The average of 40, 42, and 50 is 1. 42 2. 43 3. 44 4. 45

This section of the test is also timed. The subject is given 35 minutes to respond to 50 items.

The validity coefficients for the Numerical Section of the CQT with the first semester grade point average range from .19 to .63.

College Qualification Tests--Total Score. This score sums the scores of the Verbal, Information, and Numerical sections. The authors report that the CQT--Total Score "appears to be highly predictive of first semester grade point average." (5:45) The correlation of the CQT--Total Score with the first semester grade-point-average of beginning students ranged from a low of .34 to a high of .68 in publicly controlled four year institutions. The reliability coeffi-

cients for the total test are .97 for males and .96 for females using the odd-even coefficient of reliability.

Demographic Variables

Age. The age range of the subjects was from a low of 234 months to a high of 432 months. The mean age in months was 259.5 or just a little over 21 years and 7 months.

Sex. In the sample of 200 subjects there were 169 men and 31 women.

High School Size. Each subject was asked to report the size of the high school that he attended. This information was coded and included in the original data matrix. (See Appendix IV for the coding procedure.)

Books per year. Each subject was asked to estimate the number of books, exclusive of assigned reading and textbooks, that he read in the course of a year. The range could be from 0 to 98. If a student estimated more than 98 books per year he was coded as reading 98. 99 was reserved for indicating that no information was available. This question was included because it was thought that it might be of value to know something about the actual reading habits of the subjects.

Magazines Read Regularly. Subjects were asked to make an estimate of the number of magazines that they read regularly. Magazines read for example in the barber shop or in the dentist's office were not to be counted.

Allowable responses were from 0 to 9. If a subject indicated that he regularly read more than 9 magazines he was still coded as reading 9.

Newspapers Read Regularly. Each subject was asked to report the number of newspapers he read regularly. An occasional reading of a particular newspaper was not to be construed as regular readership. This question, like the two previous ones, was included to learn of the reading habits of the subjects. The possible range was from 0 to 9. If a subject indicated he regularly read more than 9 newspapers he was still listed as reading only 9.

College Class. Each subject indicated his class in the University: freshman, sophomore, junior, senior, graduate, and special. This question was included to learn about the composition of the sample.

Academic Major. Students were asked to indicate academic major so that the group could be described as fully as possible. As might be expected, a majority of the subjects had academic majors connected with the College of Business. A sizeable minority, however, had academic majors not connected with the College of Business. (See Table 1)

Grade Point Average. Each subject was asked to list his cumulative grade-point average so as to have one more external criterion against which to check the validity of the experimental tests. Grade-point averages were accepted

as given. (60:26)

Research Design

The research design used in this project is essentially that used in the construction of any new test. A major difficulty encountered in the making of a new test when no adequate existing test is available as a criterion measure is that of determining the validity of the new test. In order to make some estimates of the validity of the experimental tests several techniques were used. These will be discussed below.

This section of this paper discusses the validity problem, the factor analytic and correlational techniques used, the calculation of the reliability coefficients, and the construction of two indexes drawn from the factor analysis of the items of the tests.

Validity

Validity may be defined as a measure of how well a test measures whatever it is that it is supposed to measure. That is, a test that is supposed to measure how well students will do when they take a course in college algebra is valid if it allows the tester to make good predictions about student success in the mathematics course and not valid if it does not allow the tester to predict success and failure.

There are four major kinds of validity. These are

face or content validity, concurrent validity, predictive validity, and construct validity. Face validity is found by having specialists in the field being tested look at the items of the test and make a judgment as to whether or not the test seems to be measuring what it is supposed to measure. It is felt that several of the experimental tests used in this study have face or content validity.

Concurrent validity compares the test to be validated with some presently existing measure. If the test to be validated correlates highly with a presently existing test that purports to measure the same things as the new test the new test can be said to have concurrent validity.

The third kind of validity is predictive validity. This kind of validity is found by comparing the results of the test in question with some future criterion. The example given earlier of predicting success in college algebra is an example of predictive validity. Predictive validity cannot, of course, be used in this study. It is possible that given enough time, and with the selection of a suitable criterion that this kind of validation could ultimately be employed.

A technique often used today in determining test validity is factor analysis. Cronbach gives a sensible account of the way factor analysis is used when he says:

> The investigator gives a large collection of tests to the same persons. The analysis tries to determine how many distinct abilities are being measured reliably, to detect additional "trace" abilities which could be measured reliably by mod-

> ifying the tests, and to reduce the confusion which results when the same ability is given different names in different tests. Factor analysis gives information about the nature and organization of individual characteristics and clarifies what any given test measures. (25:247)

The fourth kind of validity is called construct validity. This is an analysis of the meaning of test scores in terms of psychological concepts or constructs. Construct validity is established through the interplay of observation, reasoning, and imagination. This process is much the same as that by which scientific theories are developed. Cronbach suggests that there are three parts to the establishment of construct validity. These are:

> Suggesting what constructs might account for test performance. This is an act of imagination based on observation or logical study of the test.
> Deriving testable hypotheses from the theory surrounding the construct. This is a purely logical operation.
> Carrying out an empirical study to test this hypothesis. (25:121)

Much of the work reported on in this dissertation is in the realm of construct validation.

Factor Analysis and Correlational Techniques

Several different factor analyses were done on the results of the data gathered in this study. Factor analysis was done on the items of Tests I, IV, VI, VII, and VIII. This was done as part of the item analysis and will be mentioned later.

In addition, factor analyses were performed on the total test scores of all variables. Another analysis was

done in which the demographic variables of age, sex, class, etc. were omitted and only the results of the actual tests were used. This analysis helped to make clearer what was actually being measured by the experimental tests.

The same method of factor analysis was used in each case. First, all the scores on each variable are converted into unitary standard measurements and an intercorrelation matrix computed. This matrix is then submitted to a principal axis factor analysis. Following this the principal axis analysis was rotated using a varimax procedure. A detailed discussion of the above techniques can be found in Harmon. (47)

Since the factor analysis procedure requires that there be no missing data all missing data was recoded to the mean category for the distribution. This reduces the variance of the distribution but it allows us to avoid throwing away usable data.

Item Analysis

Item analysis allows the experimenter to make judgments about the items that make up his tests. It allows him to see which items are too easy, which too difficult. It allows him to see which items correlate well with the total test score, and it allows him to see which items discriminate the high scorers from the low scorers. In the case of factor analysis of item scores it allows the experimenter to see which items are, in essence, measuring the same thing as other items.

Item analysis was done on experimental tests I, IV, VI, VII, and VIII. Tests II and III were omitted from the analysis because of their restricted range and relatively low reliability. Test V was omitted from the item analysis because it was felt that Test VIII was measuring essentially the same things more accurately.

Several kinds of item analysis were done in this study. Each of these will be discussed briefly below.

Factor Analysis of Item Scores. Factor analysis of the item scores allows the experimenter to determine what factors are being defined by the items of his test. This kind of analysis allows one to pick out of the items comprising a whole test those that have their highest loadings on certain factors and in this way increase the effectiveness of the test as a whole. This is what was done in constructing the two indexes that will be discussed below.

Item correlation with total test score. For experimental tests I, IV, VI, VII, and VIII the correlation of each item with the total score for the test was computed. This is a biserial correlation that is essentially a product moment correlation. (44:329) This correlation provides an "index of discrimination." It shows to what extent success on the item is related to success on the test as a whole. To state it another way, it tells the extent to which people who did well on the whole test did better on the item than people who did poorly on the

whole test. Diedrich (27:5) suggests that professional test constructors like to have their average biserial above .4 and are proud of themselves if it goes above .5. He also suggests that items below .3 should either be eliminated or modified.

Item ease and item difficulty. Another type of item analysis used divided the scores on Tests I, IV, VI, VII, and VIII into the high 27% of the scorers and the low 27% of the scorers. The difference between the number in the high 27% getting an item correct and the number in the low 27% getting the item correct provides a measure of item discrimination using an index developed by the Educational Testing Service. (28) The percentage of all subjects getting an item correct provides an index of item ease.

In this study the subjects on each of the tests for which the items were analyzed were divided into the high 27% and the low 27% and the proportions of success in each group on each item were computed. In addition, the total proportion of success for each item was computed. The total proportion of successes for each item is reported as p. The item difficulty score is reported as d.

Reliability

Reliability is an estimate of the correlation you would get if you administered two parallel forms of a test so closely together that no learning took place between each administration and then computed the correlation

between the scores. The reliability coefficient is, then, a measure of the "repeatability" of a test. It indicates the degree to which a test will measure a second time what it measured the first time. For a particularly clear discussion of reliability see Cronbach. (25:126-142)

Reliability coefficients for the eight experimental tests were computed using the Kuder-Richardson Formula 21. (25:141) This formula is conservative estimate of the reliability of a test. It may underestimate reliability, but it will never overestimate it. This formula was chosen because it is a conservative estimate and because it is relatively easy to compute.

The question is often asked as to what level of reliability is satisfactory. The answer has to be that reliability should be as high as possible in a given test or field within given time limits. Diederich (27:29) points out that professional test publishers are not usually satisfied with coefficients lower than .90, but he goes on to say of teacher made tests that "most of those that the writer reguarded as good, usable tests achieved reliabilities between .60 and .80."

Index Construction

Two new indexes (in effect new tests) were constructed from the items of Tests I, IV, and VI. These were constructed after examination of the factor analysis of the items of tests I, IV, VI, and VII. It was evident that

Test I had most of its items with their highest loadings on factor 2. Test VI (1-25) had most of its items loading on factor 4 along with some items from Test I and IV. Using factor 2 as the criterion for the construction of index 1, and factor 4 as the criterion for constructing index 2, 17 items drawn from test I made up index 1, and 22 items from tests I, IV, and VI made up index 2. The correlations between the indexes and all other variables were computed. This procedure allows us to take the information gained from the factor analysis and use it to make, in effect, new and more highly refined tests.

Conversion of the Iowa Silent Reading Test scores

All of the tests in the Iowa Silent Reading Tests are strictly timed. This means that those subjects who read rapidly and work rapidly have an advantage in their raw scores over those students who work and read more slowly. In order to minimize this speeding factor all the scores obtained from the Iowa Silent Reading Tests were converted into percentage scores. Both the raw scores and the percentage scores are reported.

The percentage scores were obtained by dividing the number of items attempted into the number of items right. When the quotient fell between two whole numbers the smaller was recorded as the percentage score.

The correlations between the raw scores and the percentage scores vary considerably depending upon the test.

While the correlations are in each case positive and fairly high it is clear that new information is gained by calculating the percentage score and thus minimizing the effect of the subject's rate of reading.

CH. IV. RESULTS

Introduction

This chapter discusses the results of the research under six main divisions. These are: the distribution of the scores, means, standard deviations, and standard errors of measurement; reliability; item analysis; test intercorrelations; factor analysis; and index construction.

Distribution of the Scores, Means, Standard Deviations, and Standard Errors of Measurement

All the experimental tests with the exception of Test I were scored as number right. Test I was scored using the formula: score = rights - wrongs.

Frequency polygons for each experimental test are shown in graphs 1 through 10. The graphs show that in each test the scores are skewed toward the high end of the scale. This is to be expected. The subjects were juniors and seniors and represent an intelligent and linguistically sophisticated group. They will score well on most tests of linguistic ability.

Test I has a distribution with a range of scores from 25 right down to 0 right. Tests II and III have very restricted ranges and distributions. Test II has a range of

from 40 right down to 33 right; Test III ranges from a high of 20 to a low of 14.

Test IV has a range from 25 right down to 2 right. The test is markedly skewed toward the high end of the scale.

Test V, with a possible score of 121, ranges from a high of 121 correct to a low of 5 subjects scoring between 10 and 19 correct. The graph shows a distribution that is essentially level rather than bell-shaped. The distribution shows some skewness toward the high end of the scale.

Test VI (1-25) is skewed toward the high end of the scale. Test VI (26-35) has only 10 items in it, and the shape of the distribution is reminiscent of that in Tests II and III. In Test VI (26-35), however, the range is from a low of none right (1 subject) to a high of 61 perfect scores. When Test VI is treated as a whole and incorporates all 35 items it is still skewed to the high end of the scale, but it is definitely a curve rather than a straight line.

Test VII is definitely skewed toward the high end of the scale. Five subjects had perfect scores while 27 had scores of 24 right out of 25. The distribution falls off from this peak to a low of 6 right out of 25.

Test VIII is much like Test V in the shape of its distribution. It is peaked at the high end with 27 subjects scoring either 57 or 58 right out of 60. From this peak the distribution falls off until we find one subject

scoring only 1 or 2 right out of 60.

In summary, it can be said that all of the experimental tests are skewed toward the high end of the distribution. Tests I, IV, VI (1-35), and VII have distributions that approach normality in shape. Tests II, III, V, and VIII exhibit distributions that are not normal. Tests II and III have such restricted ranges that no satisfactory judgments about the distributions can be made. Tests V and VIII, while having adequate range, have distributions that are closer to a straight line than a curve.

The mean, standard deviation, number of observations, and possible range for each of the variables measured are tabled in Table 2. Note that there were 185 or more observations for most of the variables. Exceptions are the MSU Mathematics Test with 120 observations and the MSU Arithmetic Test with only 62 observations. The MSU Arithmetic Test must not be given too much weight when it comes up in the various analyses because when the recoding is done much of the variance attributable to it will be gone.

A standard error of measurement was computed for each of the experimental tests. These are recorded in Table 3. The standard errors of measurement range from a high of 3.66 (Test V) to a low of 1.03 (Test VI, items 26-35).

Reliability

Reliability coefficients for each of the experimental

tests are reported in Table 4. These coefficients were computed using the Kuder-Richardson formula 21.

An examination of Table 4 shows that the reliability coefficients vary from a high of .98 (Test V) to a low of .31 (Test III). If we consider reliability coefficients of .70 and above to be high enough so that we can place some reliance on the tests we find that all of the tests with the exception of Test II and III are .70 or above. The last 10 items of Test VI have a reliability coefficient of .54, but when these items are combined with the first 25 of the test we find that Test VI (1-35) has a reliability coefficient of .79.

These reliability coefficients are, in general, high enough so that it is possible to place some faith in the tests. That is, Tests I, IV, V, VI, VII, and VIII all seem to measure whatever it is that they measure with some degree of reliability.

Even Test VI (26-35) can have its reliability coefficient of .54 raised by adding items to the test. The Spearman-Brown prophecy formula tells us that we can expect a reliability coefficient of .80 for this test if we increase the number of items 3.4 times. This means that Test VI (26-35) could be expected to have a reliability of .80 or above if we had 35 items rather than 10 items.

Tests II and III have relatively low reliability coefficients, and while the scores on these tests are in-

cluded in the various correlation and factor matrices the results should be reguarded with some suspicion. The reliability coefficients are so low that there is no guarantee that the tests are measuring with enough precision so that any faith can be placed in the results.

Item-Analysis

Item r with the total test score

Item correlations with the total test score are reported in Tables 5, 6, 7, 8, and 9 for the items making up Tests I, IV, VI, VII, and VIII. Tests II and III are not included in this analysis because of their low reliability. Test V was not included in this analysis because it was felt that Test VIII was measuring much the same skills as Test V. The results of the factor analysis show that this is true.

Test I has an average item r of .36. The item correlations with the total test score range from a high of .50 (items 6 and 12) to a low of .16 (item 1). If, as Diederich suggests (27:5), we should discard or reexamine items below .30, then we might wish to look closely at items 1, 8, 9, 15, 24, and 25. Even these, with the exception of item 1 range between .22 and .26.

Test IV has an average item-test correlation of .47. The highest item-test correlation is .59 (item 18), and the lowest is .14 (item 5). All the item-test correlations are reported in Table 6. If we look for items that fall

below a correlation of .30 with the total test score we find only item 5.

The correlations between the items of Test VI and the total test score are reported in Table 7. The average item-test correlation is .41; the lowest *r* is .24 (item 13); and the highest *r* is .60 (item 6). Items that have correlations with the total test below .30 are 3, 13, 14, 30, and 33. Even here we find that 3, 14, and 33 have correlations of .28 and .29.

The item-test correlations of Test VII are shown in Table 8. The average item-test *r* is .40. The highest *r* is .56 (item 7), and the lowest is .22 (item 23). Items having an item-test correlation below .30 are: 1, 10, 11, and 23.

The item-test correlations of the 60 items making up Test VIII are reported in Table 9. The average item-test correlation is .54. The lowest item-test *r* is .21 (item 2), and the highest is .74 (item 31). Items having an item-test correlation below .30 are items 1 and 2.

High 27%-Low 27% Analysis

High-low 27% item analysis was done on the tests mentioned in preceding section. The results of this analysis are reported in Tables 10, 11, 12, 13, and 14. This analysis serves as reinforcement for the findings previously reported about the items of Tests I, IV, VI, and VII. Table 15 averages the item statistics for each sentence

in Test VIII. This is done since the items making up each sentence are highly intercorrelated. Table 15 treats each sentence as a unit and allows us to see how the sentences differ in their difficulty of translation.

Factor Analysis of Items

Factor analysis gives us another way of looking at the items making up the tests. Factor analysis of the items lets us see which items measure much the same things as other items and gives us another way of looking at item reliability. A test with most of its items loading on a single factor can be considered to have more reliability than tests wherein the items of the test split among a number of factors.

The 110 items making up Tests I, IV, VI, and VII were factor analyzed using the principal axis method and a varimax rotation analysis. Eight factors accounted for 34.11% of the variance. The rotated factor loadings for the items of Tests I, IV, VI, and VII are reported in Table 20. If we look for the highest loadings for each item we find that 17 items in Test I have their highest loadings on factor 2; 3 items have their highest loading on factor 4; 1 item has its highest loading on factor 7; and 2 items have their highest loadings on factor 8. It is evident that much of the variance of Test I is accounted for by the second factor.

Test IV (the 25 item traditional part-of-speech test)

is not nearly so unifactorial as is Test I. In Test IV we find 10 items having their highest loading on factor 1; 2 on factor 3; 1 on factor 4; 7 on factor 5; 2 on factor 6; and 2 on factor 8.

Test VI (1-25), like Test I, has the majority of its 25 items with their highest loadings on a single factor. Test VI (1-25) has 16 items with their highest loadings on factor 4; 2 on factor 1; 5 on factor 5; and 2 on factor 7. Test VI (26-35) spreads its items around much more than does Test VI (1-25). Test VI (26-35) has 1 item with its highest loading on factor 2; 4 on factor 3; 2 on factor 4; 1 on factor 5; 1 on factor 7; and 1 on factor 8.

Test VII, the part-of-speech test using nonsense words, splits its items up among the factors in much the same way as does Test IV. In Test VII we find 6 items having their highest loading on factor 1; 1 on factor 3; 5 on factor 5; 5 on factor 7; and 3 on factor 8.

Table 22 shows in tabular form the number of items for each test having their highest loadings on a factor. It can be easily seen that factor 2 is largely made up of items from Test I, while factor 4 is largely made up (insofar as the high loadings go) with items from Test VI (1-25). These high item loadings associated with factors were used in making up Index 1 and Index 2 that will be discussed later.

The factor analysis of the items of Tests I, IV, VI, and VII can be summarized by saying that Tests I and VI

(1-25) are essentially unifactorial. Test IV splits its high loadings between factors 1 and 5. Test VI (26-35) spreads fairly evenly over 6 of the 8 extracted factors. Test VII also spreads itself across 6 factors, but shows some similarities in its distribution to Test IV. This is to be expected since Test VII is identical to Test IV except that the items in Test VII consist of made-up words rather than real ones.

Test Intercorrelations

Test intercorrelations are reported in Tables 16 and 17. Table 16 shows the correlations between Tests I-VIII and 20 other variables. Table 17 shows the correlations of each of the experimental tests with the other experimental tests. In any kind of correlational analysis the question of "how large must a correlation be to be significant?" must be answered. In this case we find that any correlation larger than ± .17 is significant at the .02 level. (44:207-208)

As we look at Table 16 it is possible to note some interesting results. The correlations are in general positive. Of the 180 correlations reported we find only 12 negative, and of these the largest is only -.09. This is not large enough to be significant. It is also possible to observe that Test V and Test VIII have no negative correlations with any of the other measured variables.

Of the 180 correlations reported in Table 16 we find

that 64 are in excess of the .17 that is significant at the .02 level. The variable having the largest number of significant correlations with the eight experimental tests is the sentence meaning section of the Iowa Silent Reading Test. The sentence meaning section of the Iowa has 7 out of 9 possible correlations significant. This holds true both for the raw scores on the sentence meaning section and the same scores converted to percentages. The significant correlations are with the same experimental tests with the exception of Tests II and III which reverse themselves from the raw score section to the percentage score section. This is probably not very meaningful due to the low reliability of Tests II and III.

The variables with the lowest number of significant correlations with the experimental tests are Reading Rate A and Reading Rate B from the Iowa. Neither has any r large enough to be significant. This is not surprising. None of the experimental tests were timed, and speed was not a factor.

The raw word meaning section of the Iowa and the MSU English Placement Test both had 6 correlations with the experimental tests of .17 or above. The highest correlation in the matrix is + .41. This is the correlation of the MSU English Placement Test with Test IV. Test IV is the traditional part-of-speech test using real words. The MSU English Test uses some questions like this and thus the relatively high r can be expected.

If we read the correlation matrix of Table 16 down rather than across we find that Test III has 10 significant correlations out of 20. This, again, should be treated with caution because of the low reliability of Test III. Test V has 14 significant correlations out of 20; Test VI (1-25) has 9; Tests VII and VIII each have 8; Test IV has 7; Test I has 6; and Tests II and VI (26-35) have 1 each.

In brief summary, we find that Table 16 lists no significant negative correlations and relatively few negative correlations at all. The experimental test that has the largest number of significant correlations with the 20 outside variables is Test V with 14. The next largest is Test III with 10. The experimental test with the fewest correlations with the outside variables is Test II with 1. The outside variables having the highest number of significant correlations with the experimental tests were the sentence comprehension section of the Iowa. This was true in both the raw score form and the percentage score form. The outside variables having the fewest number of significant correlations with the experimental tests were the reading rate sections and the percentage score directed reading section of the Iowa. None of these variables had any *r* that was significant.

The highest *r* is .41 with Test IV and the MSU English Placement Test. The lowest *r* is -.09 between Test II and the grade point average.

Table 17 reports the intercorrelations of the experimental tests. Of the 36 correlations, 2 are negative. Neither is large enough to be significant. Of the 34 positive correlations 25 are large enough to be considered significant at the .02 level. The largest single correlation is .60 between Test IV and Test VII. Close behind this is .49 between Test V and Test VIII and .50 between Test IV and Test VI (1-25). The smallest correlation is -.04 between tests II and IV.

Average correlations between each experimental test and all the other experimental tests are as follows: Test I, .22; Test II, .11; Test III, .15; Test IV, .26; Test V, .27; Test VI (1-25), .34; Test VI (26-35), .24; Test VII, .31; Test VIII, .28.

In summary, an examination of Table 17 shows that the experimental tests correlate, in general, positively with each other. There are only 2 negative correlations; neither is significant. The test having the largest average $\underline{r}$ with the other tests is Test VI (1-25); the test having the smallest average $\underline{r}$ with the other tests is Test II. Once again, attention should be called to the low reliability coefficients associated with Tests II and III.

Test VI (1-25) and Test VII both ask the student to deal with nonsense or made-up words. These tests have the highest average correlations with the other experimental tests.

The largest r is .60 between Tests IV and VII. This is to be expected inasmuch as the structural patterns of the sentences in both tests are the same. In Test VII the form-class words are made-up words rather than real ones. The task the subject is asked to do is the same in both tests.

The experimental tests correlate better between themselves than they do with the 20 outside variables listed in Table 16. The average intercorrelation between the experimental tests is .24, and the average intercorrelation between the experimental tests and the outside variables is .13.

Factor Analysis

The 8 Experimental Tests and 20 Outside Variables

In addition to the factor analysis of test items mentioned above, two factor analyses were done on the total scores of the tests. Table 18 lists the rotated factor loadings obtained in a 29 variable analysis containing the 20 outside variables and the 8 experimental tests. The six factors extracted account for 53% of the variance.

In this factor analysis we are looking for a factor (or factors) that can reasonably be labelled "ability to respond to the structural cues of written language." That is, we hope to find a factor that cannot be accounted for by the usual tests of reading ability, grammatical ability, or generalized verbal ability. As will be seen

below we can tentatively identify factors 5 and 6 as having some relationship to the postulated factor.

The tests associated with factor 1 (i.e., those having their highest loading on this factor) are reading rate A, reading rate B, raw directed reading, raw word meaning, and raw sentence meaning. This factor can be labelled a "speed of reading" factor. It is associated with tests where speed of reading was important. When the various tests in the Iowa Silent Reading Test are converted into percentage scores it is evident that they no longer have their highest loadings associated with factor 1.

Tests IV, VI, (1-25), VI (26-35), and VII have their highest factor loadings on factor 2. Other tests having high loadings on this factor are grade-point-average and the MSU English Placement Test. The MSU English Test, Test IV, and Test VII all test the subjects' ability to name the grammatical categories of words in sentences. That is, these tests measure the ability of the subject to name the parts-of-speech. Test VI probably also measures this ability although it goes about the testing in a way that minimizes the need to know the formal classification system of noun, verb, pronoun, etc. Grade point average actually splits between this factor and factor 3.

It seems possible to call factor 2 a "traditional grammatical factor" or an "ability to name parts-of-

speech" factor. Test I which attempts to measure the subjects' ability to respond to the "grammaticality" of utterances has its next to the highest loading on this factor. The tests loading on factor 2, then, seem to all have something to do with the ability of the subject to respond to the cues of written language in the traditional manner. Test IV is the usual part-of-speech test. Test VI also gets at the ability to categorize words in sentences but it does it in a relatively untried method. Test I asks for judgments of grammaticality, not in the sense of "good" or "bad" but in the sense of "could this sentence be a real English sentence?" Since the grade-point-average splits between this factor and factor 3, it seems that the ability to respond to tests of this kind is related to general academic success as is the ability to comprehend what one reads.

A number of tests have their highest loadings on factor 3. These include: raw comprehension, raw paragraph comprehension, percentage of comprehension, percentage of paragraph comprehension, The MSU Reading Test, CQT-V, and CQT-I.

This factor can probably safely be labelled a "reading comprehension" factor. It appears to be defined by tests that can be said to measure how well one understands what one reads, and perhaps (as evidenced by CQT-I) how much one reads.

The MSU Mathematics Test, the MSU Arithmetic Test, and

the Numerical section of the CAT are associated with factor 4. It clearly can be labelled a "numerical-mathematical" factor.

Factor 5 has percent of word meaning, percent of sentence meaning, Test II, Test III, Test V, and Test VIII loading heavily on it. If we once again disreguard Tests II and III because of their low reliability, we find factor 5 defining our ability to understand words both singly and in context, and to translate "nonsense" utterances back into real English. Factor 5 also defines our ability to react to and comprehend sentences taken as wholes. Factor 5 may be tentatively labelled as a "word knowledge plus structural meaning knowledge" factor.

Two tests have their highest loadings on factor 6. These are the percent of directed reading test and Test I. The directed reading test puts a premium on the ability of the student to quickly and accurately find the sentence in an essay that answers a particular question. Test I puts a premium on the ability of the student to determine whether a sentence could or could not be a real English sentence. Test I has a fairly high loading (its next highest) on factor 2 that we have tentatively labelled a "traditional grammatical" or "part-of-speech" factor. It would seem possible to label factor 6 an "ability to judge grammaticality" factor.

Looking at the experimental tests in terms of this factor analysis, it seems clear that the most interesting

ones are Tests I, V, and VIII. Test IV was never intended to be anything more than a traditional part-of-speech test. Test VII is the same test with "nonsense" words in place of real ones. It indicates that subjects can make the appropriate categorizations without having to have knowledge of the lexical meanings of the words. Test VI, it was hoped, would have measured much the same thing as Test IV and VII without the necessity of the subject having to actually name the part-of-speech. An examination of the factor analysis indicates that this is so.

Tests V and VIII seem to be measuring something that is not covered by any of the other experimental tests. It seems to be the skill that allows a student to respond quickly and accurately to sentences plus the ability to provide synonyms for words. Note that CQT-V has its next to the highest loading on factor 5 along with Tests V, VIII, percent of word meaning, and percent of sentence meaning.

Test I, too, is not like the others. It occupies factor 6 with the percent of directed reading test of the Iowa. It measures the ability of the subject to make judgments about the grammaticality of written sentences.

To summarize the results of this factor analysis, we can tentatively say that factor 1 is a "speed" factor; factor 2 is an "ability to name parts-of-speech" factor; factor 3 is a "generalized reading comprehension/word

comprehension" factor; factor 4 is a "numerical/mathematical" factor; factor 5 is a "word knowledge plus structural meaning knowledge" factor; and factor 6 is a "judgment of grammaticality" factor.

The Eight Experimental Tests Factored Alone

Table 21 shows the rotated factor loadings for the eight experimental tests when they are factored by themselves without reference to any of the outside variables. Three factors account for 62% of the variance.

An examination of Table 21 reveals that Test I has its highest loading on factor 3. Tests II and III load heavily on factor 2. Test IV has its highest loading on factor 1. Test V splits between factors 1 and 2 but has its highest loading on factor 2. Test VI (1-25) and Test VI (26-35) split their loading between factors 1 and 3. Test VI (1-25), however, has its highest loading on factor 1, and Test VI (26-35) has its highest loading on factor 3. Test VII has its highest loading on factor 1, while Test VIII, like Test V, splits between factor 1 and 2. Both Tests V and VIII have their highest loadings on factor 2 although they have fairly large loadings on factor 1.

Factor 1 seems to identify those tests where there is a premium placed on the ability to correctly name grammatical categories. Test IV and VII both ask the subject to name the parts-of-speech. Test VI while not asking the subject to actually name the parts-of-speech is constructed

so that a knowledge of these categories will be an aid in responding correctly. Tests IV and VII are relatively "pure" tests while Test VI does split between factors 1 and 3.

Factor 2 includes Tests II, III, V, and VIII. Although we have been wary of placing much faith in the scores of Tests II and III it is evident that they are the same kind of tests and seem to be measuring much the same things. It may be noted (44:522-523) that reliability and validity are often at cross purposes. That is, a heterogenous test may have low reliability yet have high practical validity. A homogenous test may have high reliability, but the increase in reliability may not affect the test's validity. Our present research does not allow us to judge whether the above statements apply to Tests II and III, but they provide a possible explanation for the test's high loadings on factor 2. If we ascribe some weight, albeit slight, to Tests II and III it should be recalled that both ask the subject to respond by placing words from a list in appropriate slots in an English sentence. Tests V and VIII ask the subject to make appropriate translations from "sentences" made up of words that sound somewhat like those that are required in the translation. In each of these tests we have a measure of the ability of the subject to respond when some structural constraints are placed upon the possible answers.

Factor 3 has as its most representative test, Test I. Other tests having fairly high loadings on this factor are Test VI (1-25) and Test VI (26-35). This factor can possibly be called a "judgment of grammaticality" factor.

In summary, we find that the eight experimental tests can be described in terms of three factors. The three factors can be identified as: 1, a "part-of-speech" or "ability to name grammatical categories" factor; 2, an "ability to respond when certain structural constraints are placed on the response" factor; and 3, a "judgment of grammaticality" factor.

Construction of Indexes

In an effort to construct purer tests two indexes were constructed on the basis of the information in the factor analysis of the items of Tests I, IV, VI, and VII. Table 22 shows that Test I had 17 of its 25 items with their highest loading on factor 2 of the item factor analysis. These 17 items were used in constructing Index 1. The same table shows that Test VI (1-25) had 16 of its 25 items loading highest on factor 4. These 16 items plus the 3 items from Test I, 1 item from Test IV, and 2 items from Test VI (26-35) with their highest loadings on factor 4 were used in constructing Index 2. Index 1 might be called a "judgment of grammaticality index" and Index 2 may be called a "judgment of similar structural categories index."

After the two indexes were constructed the correlations

between the indexes and all the variables were obtained. The results were not so useful as had been expected. An examination of Table 19 (listing the correlations between the two indexes and the other variables) shows that the correlation between Index 1 and Test I, from which most of the index items were drawn, is .85. If the index had been of much help in refining the test we would expect that the correlation would be considerably lower. The correlation between Index 2 and Test VI (1-25), from which most of the Index 2 items were drawn, shows an even higher correlation of .93.

If one looks at the correlations between Index 1 and the outside variables and compares these correlations to those between Test I and the same outside variables it is immediately evident that they are, in general, about the same. While there are some changes they are relatively minor. The same holds true in comparing the correlations between Index 2 and the outside variables and the correlations between Test VI (1-25) and the outside variables.

It can probably be assumed that Tests I and Test VI (1-25) are relatively "pure" as they stand and that the construction of the indexes does not add a significant amount of additional information.

Summary

We find that six of our eight experimental tests are reliable according to practical standards. Two are not. The standard errors of measurement are fairly small and in line with expected figures. (27:14) An examination of the frequency polygons shows that each distribution is skewed to the high end of the scale. Item analysis indicates that the majority of items in Tests I, IV, VI, and VII are within acceptable limits in regard to both discrimination and difficulty.

Test intercorrelations are, in general, what might be expected in testing verbal ability. That is, they are in the main positive; and like tests seem to go together. The experimental tests correlate better between themselves than they do with the outside variables.

Factor analysis of the experimental tests and the outside variables identifies six factors. The factors can be labelled: 1. speed of reading factor; 2. ability to name the parts-of-speech factor; 3. general reading comprehension factor; 4. numerical-mathematical factor; 5. word knowledge plus structural meaning knowledge factor; 6. ability to make judgments of grammaticality factor. The two factors of most interest for the purposes of the present research are factors five and six.

When the eight experimental tests are factored separately three factors are isolated. Factor one identifies

the ability to correctly name grammatical categories; factor two identifies the ability of the subject to respond when certain structural constraints are placed upon the possible answer; and factor three is an ability to make judgments about the grammaticality of written utterances. The factors of interest are two and three, corresponding to factors five and six in the larger analysis.

The indexes were constructed in an effort to sharpen and intensify the effects of Tests I and VI. The results were disappointing. The correlations between the indexes and the other variables remained about the same as the correlations between Tests I and VI and the other variables.

Tests I, V, VI, and VIII, defining factors five and six in the 29 variable analysis, and factors two and three in the 9 variable analysis are promising in that they point the way to further research in this area.

CH. V. DISCUSSION

Conclusions

Two major questions must be raised about any test. These are: (1) Does the test measure reliably whatever it is that it measures? (2) What does the test measure? The first question deals with reliability; the second with validity.

Reliability can be assessed in a number of ways. The method used in this study was that of internal consistency as estimated by the Kuder-Richardson formula 21. It can be safely claimed that six of the eight experimental tests have reliability coefficients high enough so that they can be trusted. The coefficients would not satisfy a commercial test publishing company, but they are high enough for experimental purposes. Tests II and III have reliability coefficients so low that the scores on these tests cannot be trusted. The technique, however, that is used in Tests II and III seems to show some promise and might profitably be tried again using a group of subjects that are younger and less sophisticated linguistically.

All the tests with the exception of Tests II and III have reliability coefficients in excess of .70. Tests V

and VIII have reliability coefficients of .98 and .94. These are adequate by any reasonable standard.

The question of validity is another matter. In a study of this type where there is no known test that attempts to directly measure the ability to respond to the structural cues of written language, the problem is in finding a satisfactory criterion. Since we have several tests that purport to measure something that has not been previously measured (at least directly) we hope by using factor analysis to find our experimental tests loading essentially alone on one or two factors and with the outside tests not loading on these factors or loading on them to a lesser degree than the experimental tests. As we look at Table 18 we find that Tests I, V, and VIII approach this. While the factors of interest (5 and 6) are by no means pure it is evident that Tests V and VIII have their highest loadings on factor five and that their next to the highest loadings are considerably lower than their highest. Other tests having high loadings on factor five are percentage of word meaning, percentage of sentence meaning, and the verbal section of the CQT. While the factor is not as clear-cut as we might hope for it certainly shows some promise. The tests loading on factor five seem to be measuring some combination of word knowledge and sentence meaning knowledge. The word meaning knowledge is, of course, based on knowledge of lexical items. The sentence

meaning is based on the ability to respond quickly and accurately to sentences as a whole. The validity of Tests V and VIII is then related to the ability to respond not only to words alone, but also to sentences.

Test I asked the subjects to make judgments about the possible grammaticality of twenty-five utterances. Note that these judgments of grammaticality were not based on "good" or "bad" grammar in the prescriptive sense, but on whether the utterance could be or could not be a real English utterance. That is, did it or did it not have the structure of English? This test splits its high loadings between factors 2 and 6. Factor 2 can be easily identified as being the ability to name the parts-of-speech. Factor 6 is less obvious. We have tentatively identified it as "ability to make judgments of grammaticality." The only outside test having a high loading on this factor is the percent of directed reading test of the Iowa. This test (directed reading) asks the subject to quickly find the sentence in an essay that answers a question about the facts in the essay. It seems reasonable to assume that the ability to react quickly to the sentences in the essay --to be able to say accurately "what is this sentence about?" should be positively related to the question "can this utterance be a real English sentence?"

We can sum up the experimental tests then as follows: (1) Test I has adequate reliability and some face and

factorial validity. (2) Test II and Test III have such low reliability that no real effort was made to attempt to find what kind of validity they had. (3) Tests IV, VI, and VII have adequate reliability and seem to be measuring simply the ability to name the parts-of-speech. Test VI shows some promise for further development inasmuch as it provides us with a way of testing the grammatical knowledge of a subject who has had no instruction in the formal classification system of the language. (4) Test V and VIII have high reliability. They appear to have some face validity in that it seems impossible to make the required translations simply on a knowledge of word meanings alone. The translations must be made partially on the basis of word knowledge and partially on the basis of knowledge of the structure of the language. The factor analysis shows that both skills are contained in the test. This technique of testing shows some promise for the future.

Some generalizations about the research seem possible. The research indicates that there are ways to measure verbal abilities that have been little investigated. Also, it seems possible to conclude that college juniors and seniors are by no means uniform in their ability to respond to tests purporting to measure structural awareness. The research is frankly exploratory in nature and fraught with all the dangers attendant upon investigating new areas. The possible rewards are great however, for studies of this

type may have much to teach us about how we read and how we react to written language.

Implications for Further Research

There are a number of ways in which additional research might be conducted. The various tests can be refined by deleting items that do not come up to standard in regard to the item analysis; the tests can be made longer in order to provide additional reliability; modifications can be made in the form of the tests. Tests II and III might profitably be redone to see if they could be made more reliable and more discriminatory.

Certain additional statistical techniques might provide additional insight. Multiple and partial correlational techniques could prove to be valuable in determining the predictive abilities of the tests or combinations of tests.

Certainly the tests should be administered to more subjects and to different kinds of subjects. Normalization data should be collected at various age levels and educational levels. Perhaps if the subjects were students at a lower educational level the skewness of the distributions would not be so marked.

Additional outside criteria might be included. More and different reading tests might prove valuable. A search might be made for other kinds of tests that might provide additional information about the validity of the experimental tests.

New testing techiques might be sought for that would further improve our ability to determine the subject's ability to respond to structural cues. Perhaps the testing techniques could be adapted to the spoken rather than the written language.

While the present study has been largely confined to the responses that subjects made to the written language, it should be possible to look at the stimulus as well as the response. That is, which sentences are the easiest to translate? Which are the most difficult? What is the nature of an "ungrammatical" utterance? Which word-classes are the most difficult for subjects to put back into the blank spaces as in Tests II and III? Further investigations might be made into the nature of the ability to react to the nonsense words as in Test VII. We know that the correlation between the scores on Tests IV and VII (asking, in effect, identical questions) is high---why isn't it perfect? We might profitably look for the deviant responders and treat them as a unit. It should be possible to devise tests that ask subjects to make judgments about the immediate constituent structure of the written utterance and use this as a way of testing the ability to respond to structural cues.

There is certainly much research to be done. This study demonstrates that it is possible to test language abilities using new techniques, and that these tests have reliability and validity. With research like this we are

on our way to being able to systematically investigate the relationships which hold between the structure of a language and, perhaps, the lexicon and the semantics of the language.

SELECTED BIBLIOGRAPHY

1. Anderson, Wallace L., and Norman C. Stageberg (Eds.) Introductory Readings on Language. New York: Holt, Rinehart and Winston, Inc., 1962.

2. Artley, A. S. The Appraisal of Reading Comprehension. The Journal of Educational Psychology. 1943, 34, 55-60.

3. Bach, Emmon. An Introduction to Transformational Grammars. New York: Holt, Rinehart and Winston, Inc., 1964.

4. Barnette, W. Leslie, Jr. (Ed.) Readings in Psychological Tests and Measurements. Homewood, Illinois: The Dorsey Press, Inc., 1964.

5. Bennett, George K., Marjorie G. Bennett, Winburn L. Wallace, and Alexander G. Wesman. College Qualification Tests. Manual. New York: The Psychological Corporation, 1961.

6. Berlo, David K. The Process of Communication. New York: Holt, Rinehart and Winston, Inc., 1960.

7. Bloch, Bernard, and George L. Trager. Outline of Linguistic Analysis. Baltimore: Linguistic Society of America, 1942.

8. Bloomfield, Morton W., and Leonard Newmark. A Linguistic Introduction to the History of English. New York: Alfred A. Knopf, 1963.

9. Bloomfield, Leonard. Language. New York: Henry Holt and Company, 1933.

10. Brown, Dona Worral, Wallace C. Brown, and Dudley Bailey. Form in Modern English. New York: Oxford University Press, 1958.

11. Brown, Roger W. Linguistic Determinism and the Part of Speech. Journal of Abnormal and Social Psychology. 1957, 55, 1-5.

12. Brown, Roger W., Words and Things, Glencoe, Illinois: The Free Press, 1958.

13. Brown, Roger W., and Ursula Bellugi. Three Processes in the Child's Acquisition of Syntax. Harvard Educational Review. 1964, 34, 133-151.

14. Bruner, Jerome S., Jacqueline J. Goodnow, and George A. Austin. A Study of Thinking. New York: Science Editions, Inc., 1962.

15. Buchanan, Cynthia D. A Programmed Introduction to Linguistics. Boston: D.C.Heath and Company, 1963.

16. Carroll, John B. A Factor Analysis of Verbal Abilities. Psychometrika. 1941, 6, 279-307.

17. Carroll, John B. The Analysis of Verbal Behavior. Psychological Review. 1944, 51, 102-119.

18. Carroll, John B. The Study of Language. Cambridge, Massachusetts: Harvard University Press, 1955.

19. Chace, Howard L. Anguish Languish. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1956.

20. Cherry, Colin. On Human Communication. New York: Science Editions, Inc., 1961.

21. Chomsky, Noam. Syntactic Structures. The Hague: Mouton and Co., 1957.

22. Clark, Ruth Anne. The Influence of Syntactic and Referential Information on Word Predictability. Unpublished paper read before the Behavioral Sciences Interest Group of the Speech Association of America, December, 1964.

23. Cofer, Charles N. and Barbara S. Musgrave (Eds.) Verbal Learning and Verbal Behavior. New York: McGraw-Hill Book Company, Inc., 1961.

24. Cofer, Charles N. and Barbara S. Musgrave (Eds.) Verbal Behavior and Learning. New York: McGraw-Hill Book Company, 1963.

25. Cronbach, Lee J. Essentials of Psychological Testing. New York: Harper and Row, 1949.

26. Davis, Frederick B. Fundamental Factors of Comprehension in Reading. Psychometrika. 1944, 9, 185-197.

27. Diederich, Paul B. Short-Cut Statistics for Teacher-Made Tests. Princeton, New Jersey: Educational Testing Service, 1960.

28. Fan, Chung-Teh. Item Analysis Table. Princeton, New Jersey: Educational Testing Service, 1952.

29. Fiedler, Leslie. The Art of the Essay. New York: Thomas Y. Crowell Co., 1958.

30. Francis, W. Nelson. The Structure of American English. New York: The Ronald Press Co., 1958.

31. Fries, Charles C. American English Grammar. New York: Appleton-Century-Crofts, Inc., 1940.

32. Fries, Charles C. Teaching and Learning English as A Foreign Language. Ann Arbor, Michigan: University of Michigan Press, 1946.

33. Fries, Charles C. The Teaching of English. Ann Arbor, Michigan: The George Wahr Publishing Co., 1949.

34. Fries, Charles C. The Structure of English. New York: Harcourt, Brace and Company, 1952.

35. Fries, Charles C. Linguistics and Reading. New York: Holt, Rinehart and Winston, Inc., 1963.

36. Fruchter, Benjamin. The Nature of Verbal Fluency. Educational and Psychological Measurement. 1948, 8, 33-47.

37. Garrett, Henry E. Statistics in Psychology and Education. New York: Longmans, Green, and Co., 1958.

38. Gleason, H.A. Jr. An Introduction to Descriptive Linguistics. New York: Holt, Rinehart and Winston, 1961.

39. Gray, Giles W. and Claude M. Wise. The Bases of Speech: Third Edition. New York: Harper and Brothers, Publishers, 1959.

40. Gray, W. S. and Bernice E. Leary. What Makes A Book Readable. Chicago: University of Chicago Press, 1935.

41. Greenberg, Joseph H. Essays in Linguistics. Chicago: University of Chicago Press, 1957.

42. Greene, H. A., A. N. Jorgensen, and V. H. Kelley. Iowa Silent Reading Tests: New Edition. New York: World Book Company, 1943.

43. Greene, H. A., A. N. Jorgensen, and V. H. Kelley. Manual of Directions: Iowa Silent Reading Tests. New York: World Book Company, 1943.

44. Guilford, J. P. Fundamental Statistics in Psychology and Education, 2nd Edition. New York: McGraw-Hill Book Company, 1950.

45. Guth, Hans P. English Today and Tomorrow: A Guide for Teachers of English. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1964.

46. Guyer, Byron, and Donald A. Bird. Patterns of Thinking and Writing. San Francisco: Wadsworth Publishing Co., 1959.

47. Harman, Harry H. Modern Factor Analysis. Chicago: The University of Chicago Press, 1960.

48. Harris, Chester W. An Exploration of Language Skill Patterns. The Journal of Educational Psychology. 1948, 39, 321-336.

49. Harris, Zellig S. Methods in Structural Linguistics. Chicago: University of Chicago Press, 1951.

50. Hill, Archibald A. Introduction to Linguistic Structures. New York: Harcourt, Brace and Company, 1958.

51. Hill, Archibald A. Grammaticality. Word. 1961, 17, 1-10.

52. Hockett, Charles F. A Course in Modern Linguistics. New York: The Macmillan Company, 1958.

53. Hughes, John P. The Science of Language. New York: Random House, 1962.

54. Hunnicutt, C. W. and William J. Iverson (Eds.) Research in the Three R's. New York: Harper and Brothers, 1958.

55. Ives, Sumner. A New Handbook for Writers. New York: Alfred A. Knopf, Inc., 1960.

56. Jesperson, Otto. Growth and Structure of the English Language. Garden City, New York: Doubleday and Company, 1955.

57. Johnson, Donald M. and Floyd Reynolds. A Factor Analysis of Verbal Ability. The Psychological Record. 1941, 4, 183-195.

58. Joos, Martin (Ed.) Readings in Linguistics. Washington, D.C.: American Council of Learned Societies, 1957.

59. Kelley, T. L. Essential Traits of Mental Life. Harvard Studies in Education. 1935, 26, 146pp.

60. Krumboltz, John. A Note on the Accuracy of Self-Reported Grade Point Averages. Michigan State University College of Education Quarterly. Winter, 1960, p. 26.

61. Lado, Robert. Linguistics Across Cultures. Ann Arbor, Michigan: The University of Michigan Press, 1957.

62. Lado, Robert. Language Testing. London; Longmans, Green and Co., Ltd., 1961.

63. Laird, Charlton. The Miracle of Language. New York: Fawcett Publications, Inc., 1957.

64. Laird, Charlton. Thinking About Language. New York: Holt, Rinehart and Winston, 1960.

65. Langsam, Rosalind S. A Factorial Analysis of Reading Ability. Journal of Experimental Education. 1941, 10, 57-63.

66. LeFevre, Carl A. Linguistics and the Teaching of Reading. New York: McGraw-Hill Book Company, 1964.

67. Leopold, Werner F. Bibliography of Child Language. Evan-Illinois: Northwestern University Press, 1952.

68. Lloyd, Donald J. and Harry W. Warfel. American English in Its Cultural Setting. New York: Alfred A. Knopf, Inc., 1956.

69. Long, Ralph B. The Sentence and Its Parts: A Grammar of Contemporary English. Chicago: The University of Chicago Press, 1961.

70. Lunt, Horace G. (Ed.) Proceedings of the Ninth International Congress of Linguists, Cambridge, Massachusetts, August 21-31, 1962. The Hague: Mouton and Co., 1964.

71. Maclay, Howard, and Mary D. Sleator. Responses to Language: Judgements of Grammaticalness. International Journal of American Linguistics. 1960, 26, 275-282.

72. Martinet, Andre. Elements of General Linguistics. Chicago; University of Chicago Press, 1960.

73. Miller, George A. and Jennifer A. Selfridge. Verbal Context and the Recall of Meaningful Material. The American Journal of Psychology. 1950, 63, 176-185.

74. Miller, George A. Language and Communication. New York: McGraw-Hill Book Company, Inc., 1951.

75. Miller, George A., Eugene Galanter, and Karl H. Pribam. Plans and the Structure of Behavior. New York: Henry Holt and Company, 1960.

76. Nida, Eugene A. A Synopsis of English Syntax. Norman, Oklahoma: Summer Institute of Linguistics, 1960.

77. Noble, Clyde E. An Analysis of Meaning. Psychological Review. 1952, 59, 421-430.

78. Pooley, Robert C. Teaching English Grammar. New York: Appleton-Century-Crofts, Inc., 1957.

79. Potter, Simeon. Language in the Modern World. Baltimore: Penguin Books, 1960.

80. Quirk, Randolph. The Use of English. New York: St. Martin's Press, 1962.

81. Roberts, Paul. Patterns of English. New York: Harcourt, Brace and Company, 1956.

82. Roberts, Paul. Understanding English. New York: Harper and Brothers, 1958.

83. Roberts, Paul. English Sentences. New York: Harcourt, Brace and World, Inc., 1962.

84. Roberts, Paul. English Syntax. New York: Harcourt, Brace and World, Inc., 1964.

85. Robertson, Stuart and Frederic G. Cassidy. The Development of Modern English: Second Edition. New York: Prentice-Hall, Inc., 1954.

86. Rogers, Cyril A. The Structure of Verbal Fluency. British Journal of Psychology. 1953, 44, 368-380.

87. Rycenga, John A. and Joseph Schwartz (Eds.) Perspectives on Language. New York: The Ronald Press Company, 1963.

88. Sapir, Edward. Language: An Introduction to the Study of Speech. New York: Harcourt, Brace and Company, 1921.

89. Sapir, Edward. Culture, Language, and Personality: Selected Essays. (Edited by David G. Mandelbaum). Berkeley, California: University of California Press, 1957.

90. Saporta, Sol (Ed.) Psycholinguistics: A Book of Readings. New York: Holt, Rinehart and Winston, 1961.

91. Schlauch, Margaret. The Gift of Language. New York: Dover Publications, Inc., 1942.

92. Shannon, Claude E. and Warren Weaver. The Mathematical Theory of Communication. Urbana, Illinois: University of Illinois Press, 1949.

93. Sledd, James. A Short Introduction to English Grammar. Chicago; Scott, Foresman and Company, 1959.

94. Smith, Henry Lee. Linguistic Science and the Teaching of English. Cambridge, Massachusetts: Harvard University Press, 1956.

95. Sorensen, Holger Sten. Word Classes in Modern English with Special Reference to Proper Names with an Introductory Theory of Grammar, Meaning and Reference. Copenhagen: G.E.C.Gad, Publisher, 1958.

96. Stolurow, Lawrence M. and J. Robert Newman. A Factional Analysis of Objective Features of Printed Language Presumably Related to Reading Difficulty. Journal of Educational Research. 1959, 52, 243-251.

97. Sturtevant, Edgar H. An Introduction to Linguistic Science. New Haven, Connecticut: Yale University Press, 1947.

98. Taylor, Wilson L. "Cloze" readability scores as indices of individual differences in comprehension and attitude. Journal of Applied Psychology. 1957, 41, 19-26.

99. Testing Bulletin No. 3. The Use of Orientation Test Data. Office of Evaluation Services, University College, Michigan State University, July, 1960.

100. Thorndike, Robert L. and Elizabeth Hagen. Measurement and Evaluation in Psychology and Education. New York: John Wiley and Sons, Inc., 1955.

101. Thurstone, L. L. Multiple-Factor Analysis. Chicago: University of Chicago Press, 1947.

102. Trager, George L. and Henry Lee Smith, Jr. An Outline of English Structure. Washington, D.C: American Council of Learned Societies, 1957.

103. Vallins, G.H. The Pattern of English. Baltimore, Maryland: Penguin Books, Inc., 1957.

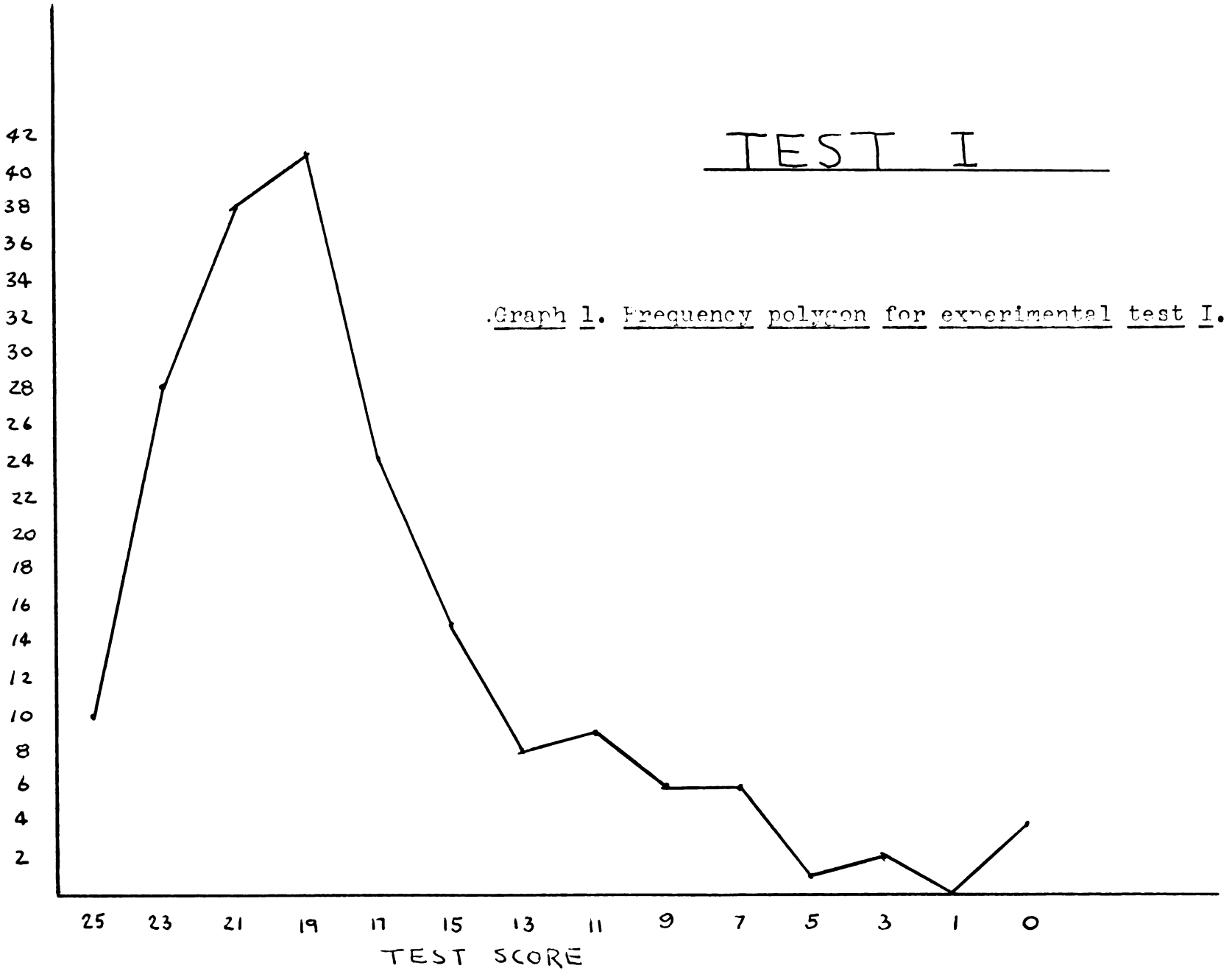
104. Vernon, Phillip E. The Structure of Human Abilities. London: Methuen and Co., Ltd., 1961.

105. Waterman, John T. Perspectives in Linguistics. Chicago: University of Chicago Press, 1963.

106. Whatmough, Joshua. Language: A Modern Synthesis. New York: New American Library, 1956.

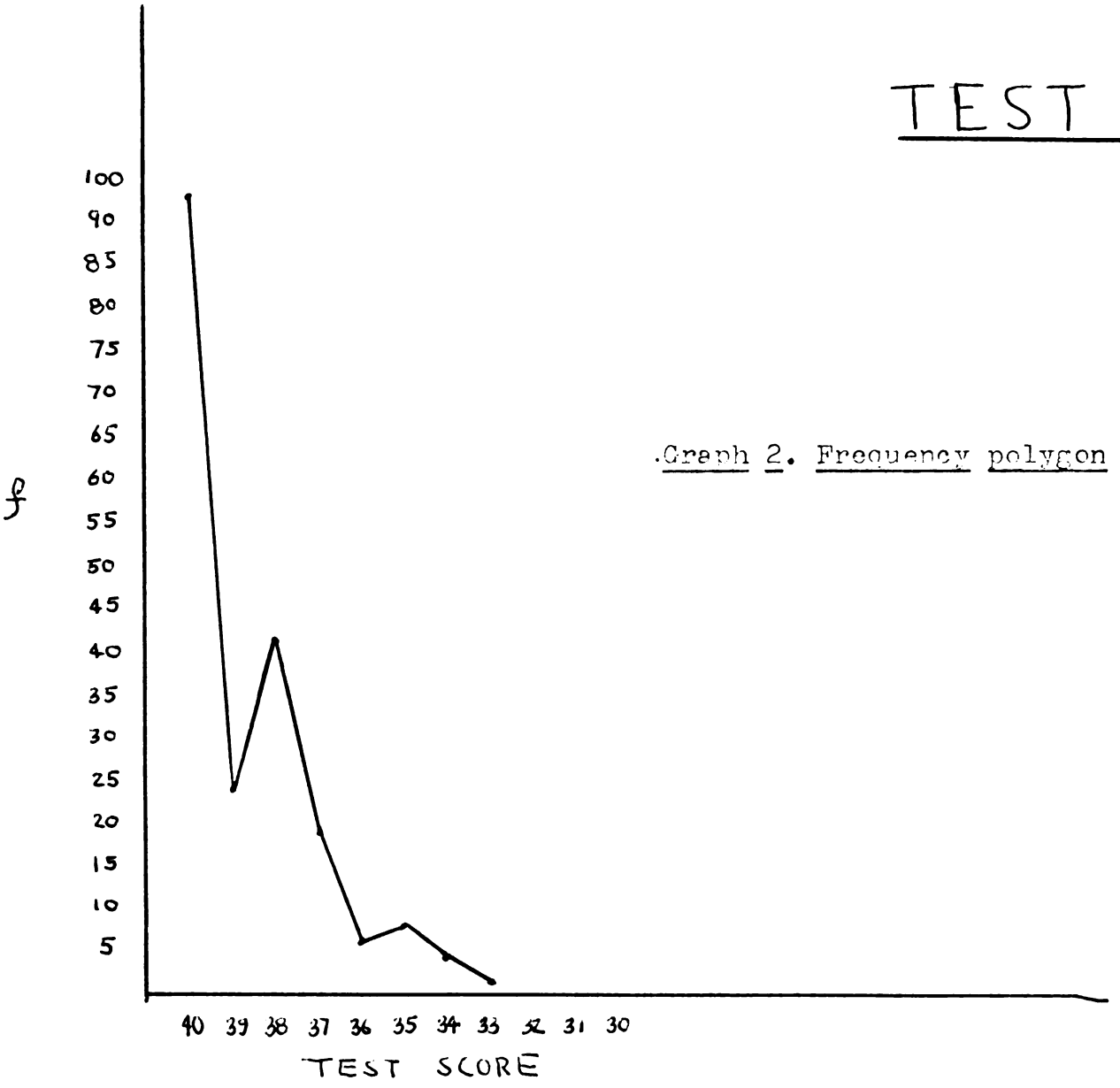
107. Whitehall, Harold. Structural Essentials of English. New York: Harcourt, Brace and Company, 1956.

108. Whorf, Benjamin Lee. Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf (Edited by John Carroll). New York: John Wiley and Sons, Inc., 1956.



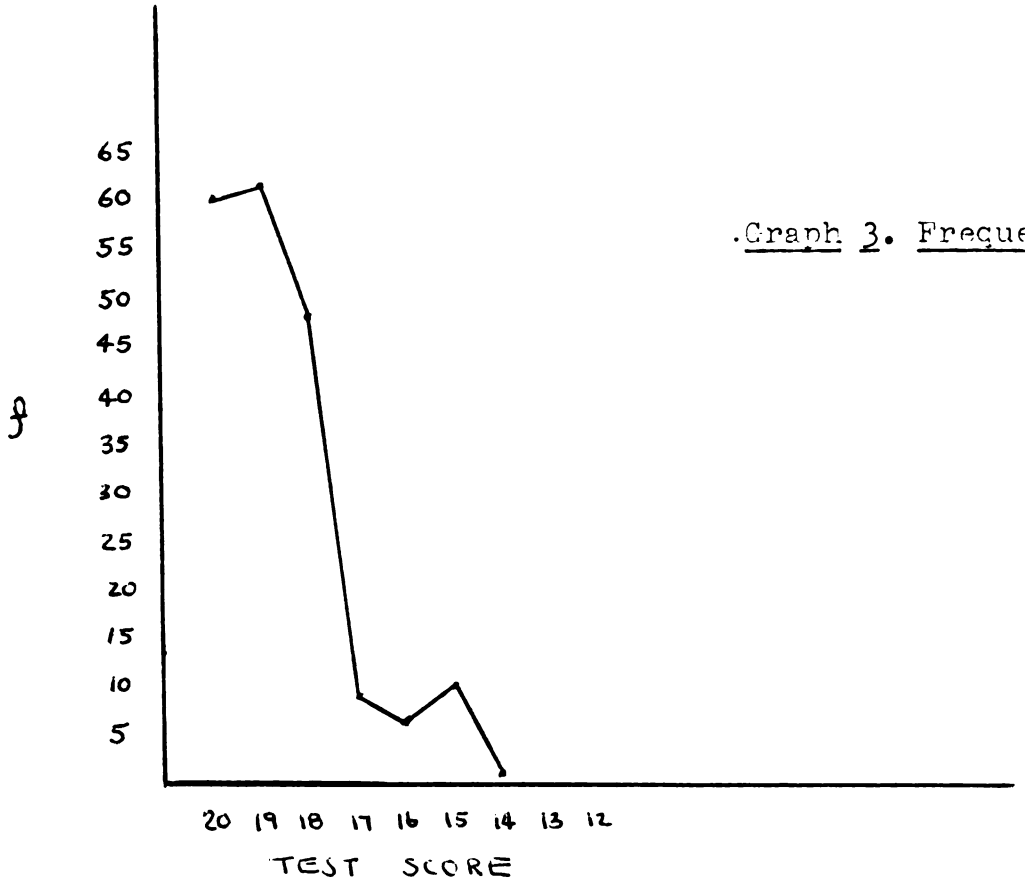
Graph 1. Frequency polygon for experimental test I.

TEST II

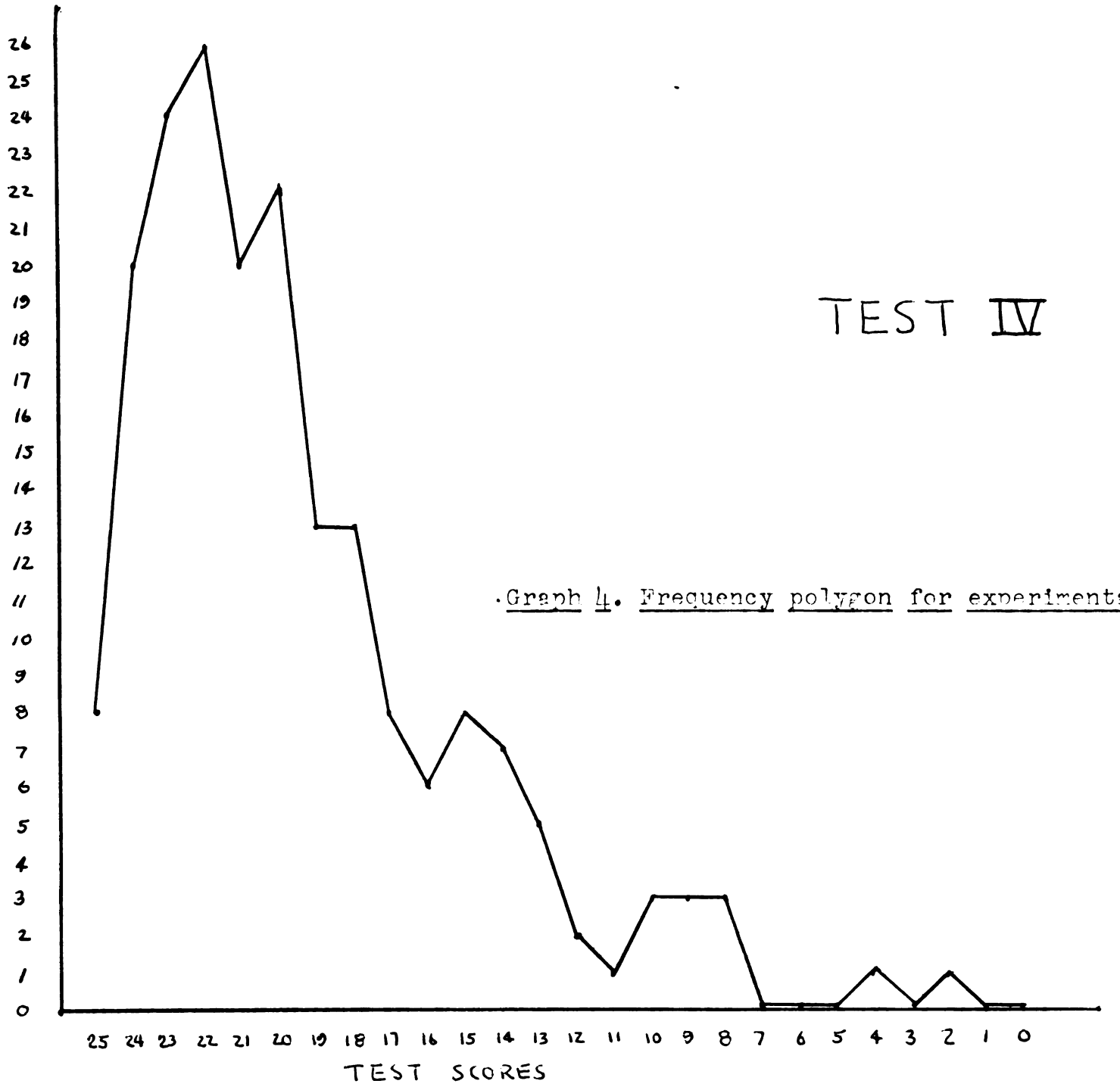


Graph 2. Frequency polygon for experimental test II.

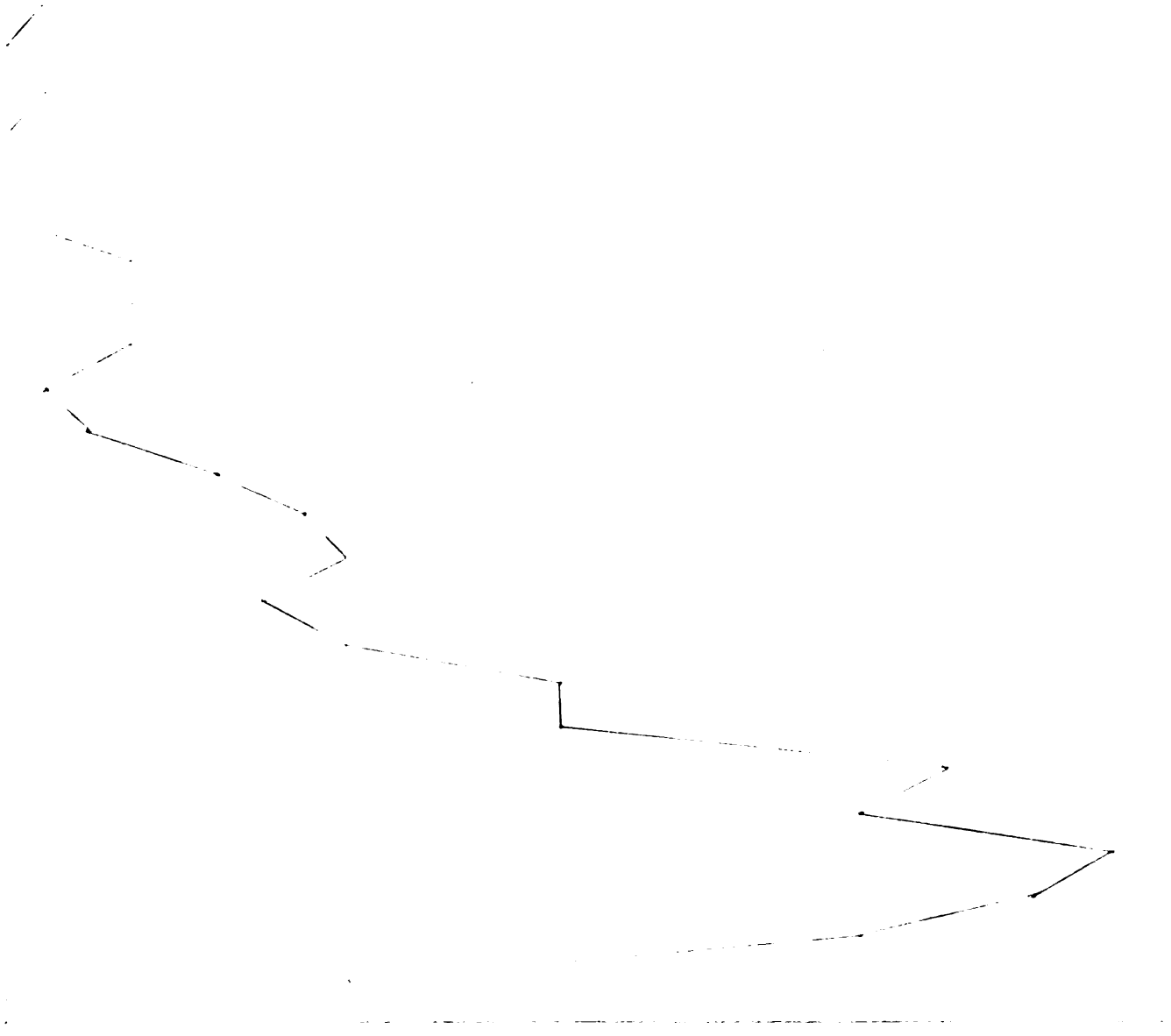
TEST III



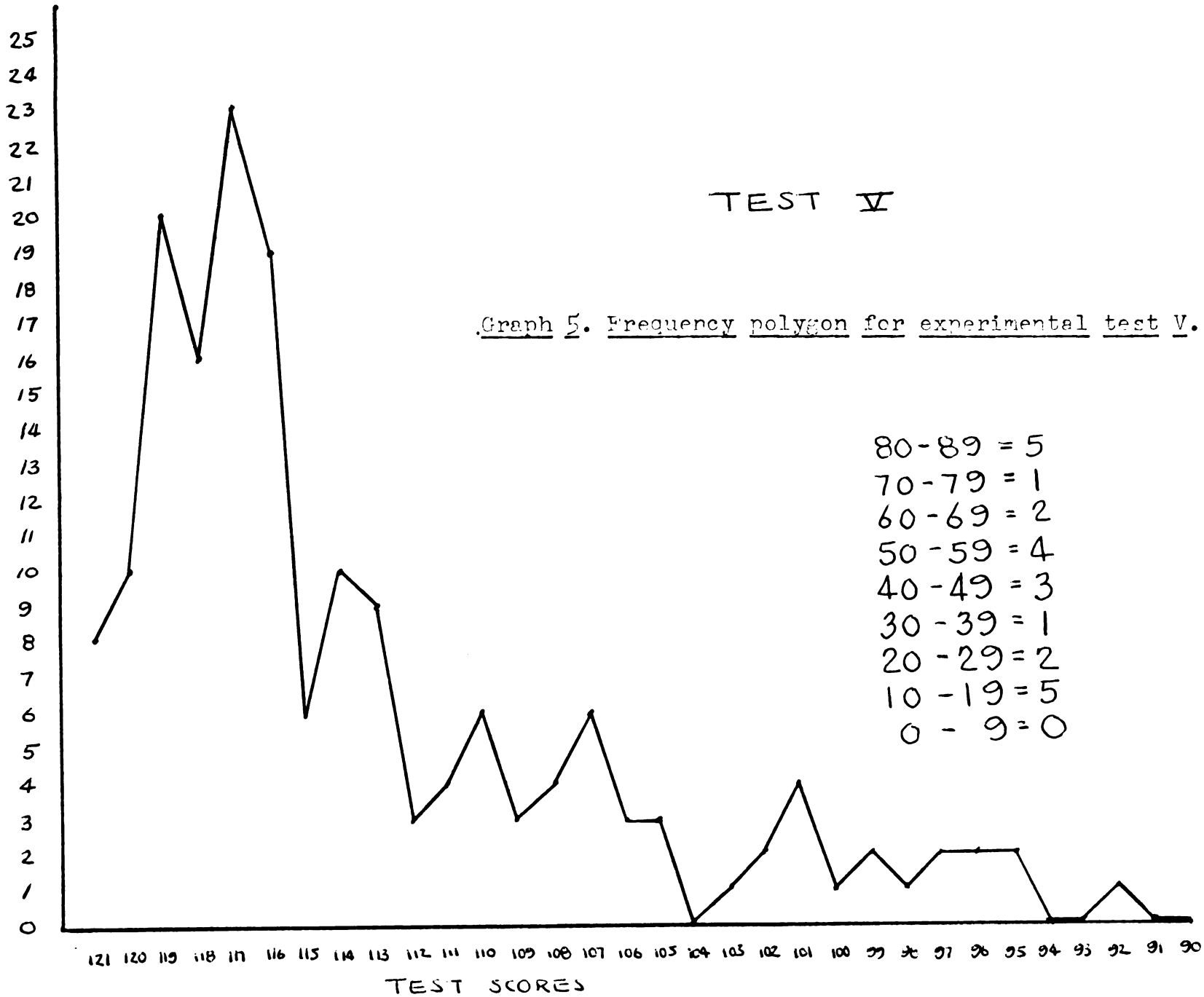
Graph 3. Frequency polygon for experimental test III.



Graph 4. Frequency polygon for experimental test IV.

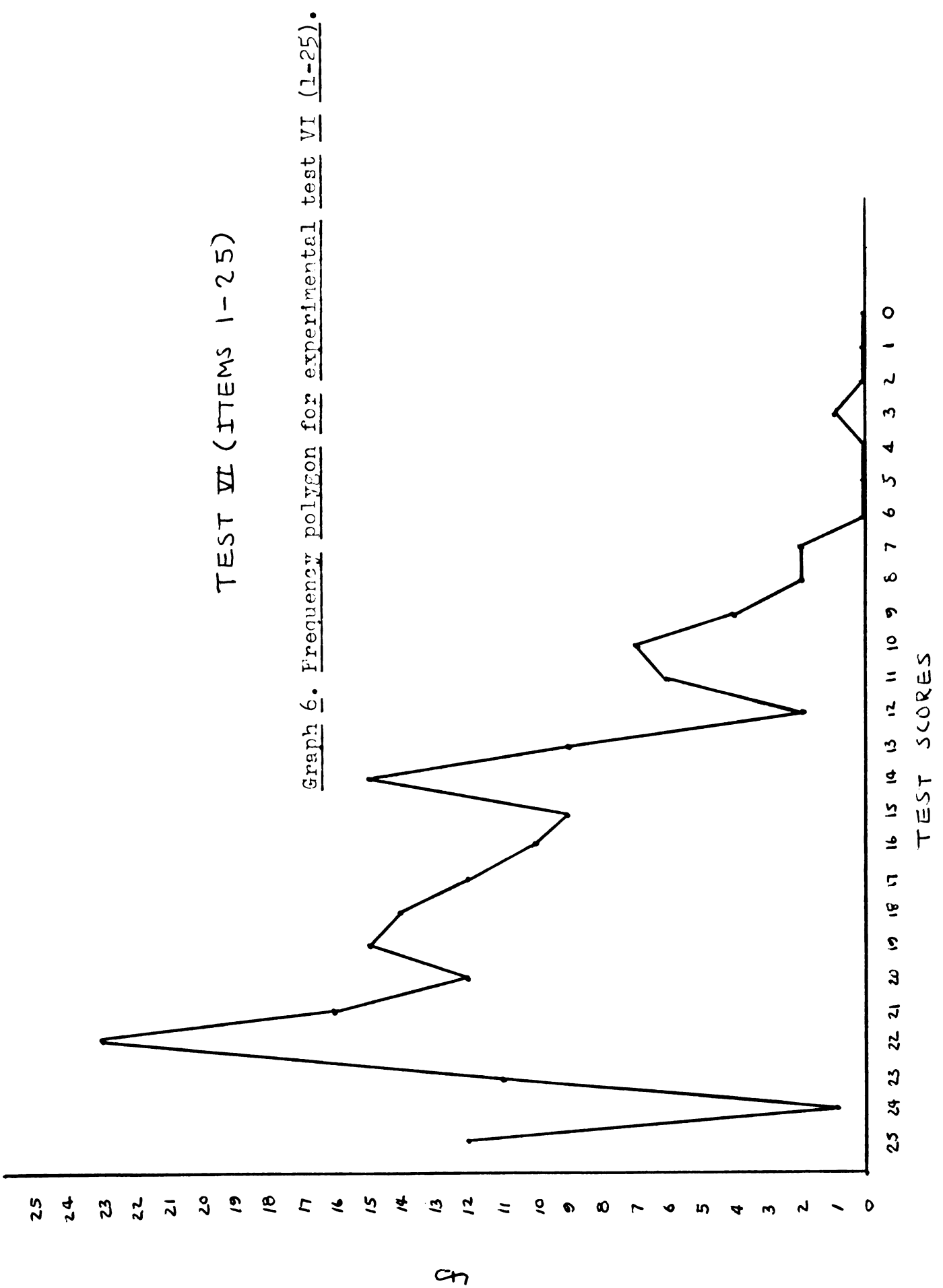


TEST V



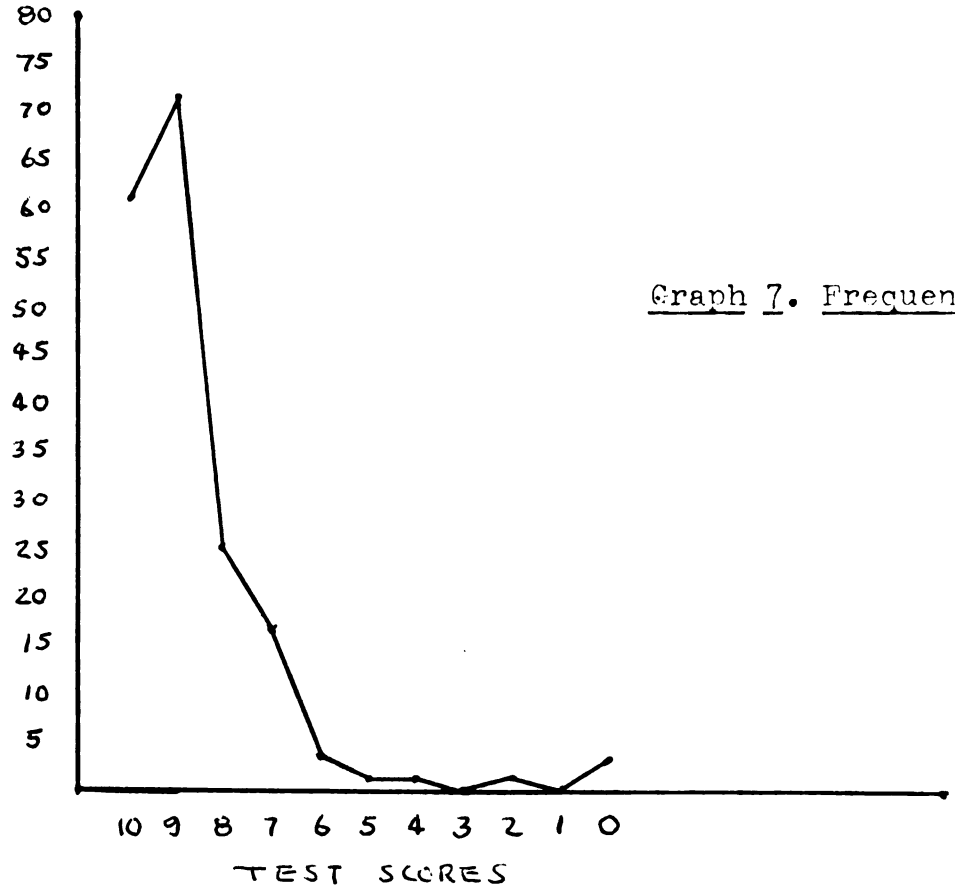
Graph 5. Frequency polygon for experimental test V.

80-89 = 5
70-79 = 1
60-69 = 2
50-59 = 4
40-49 = 3
30-39 = 1
20-29 = 2
10-19 = 5
0-9 = 0

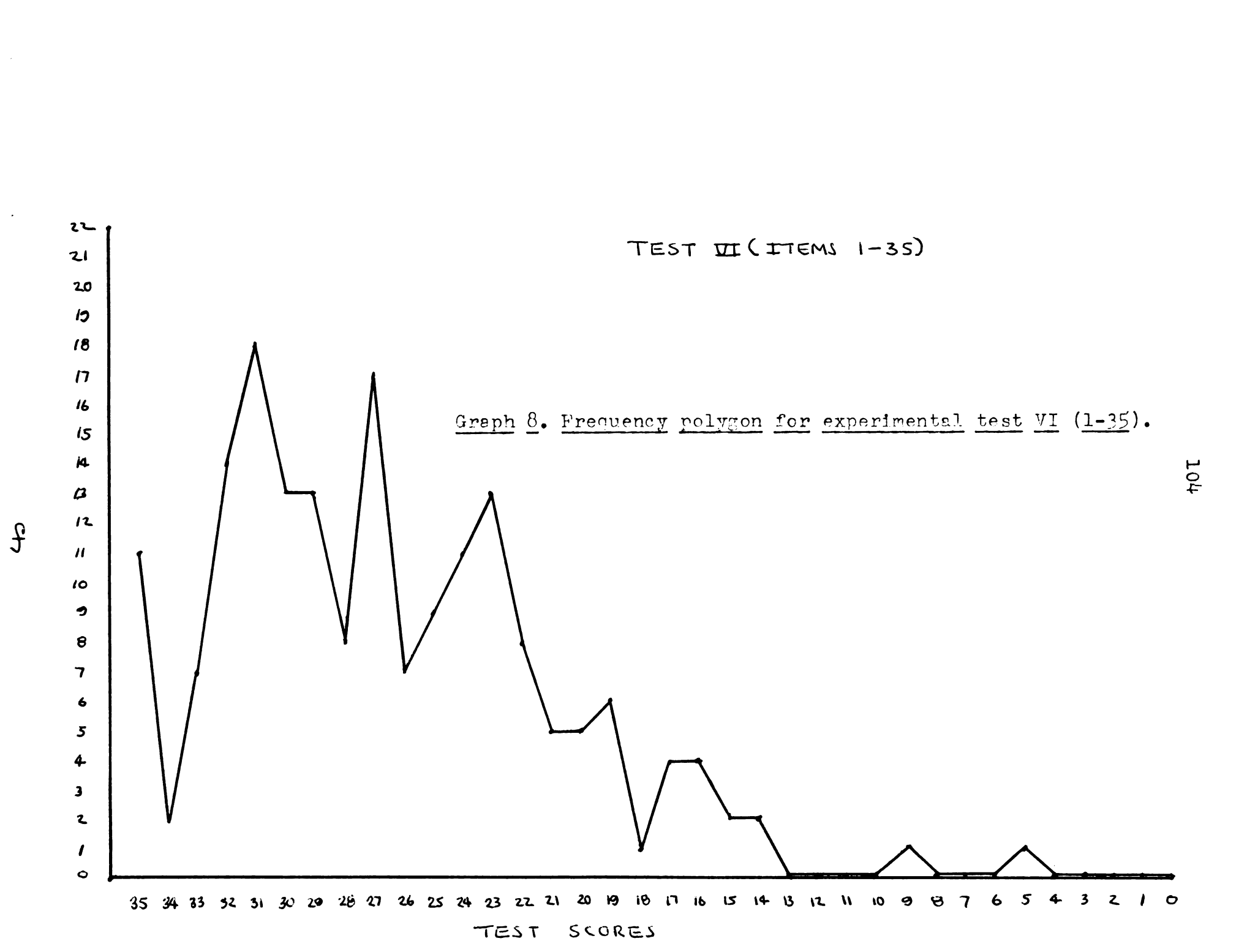


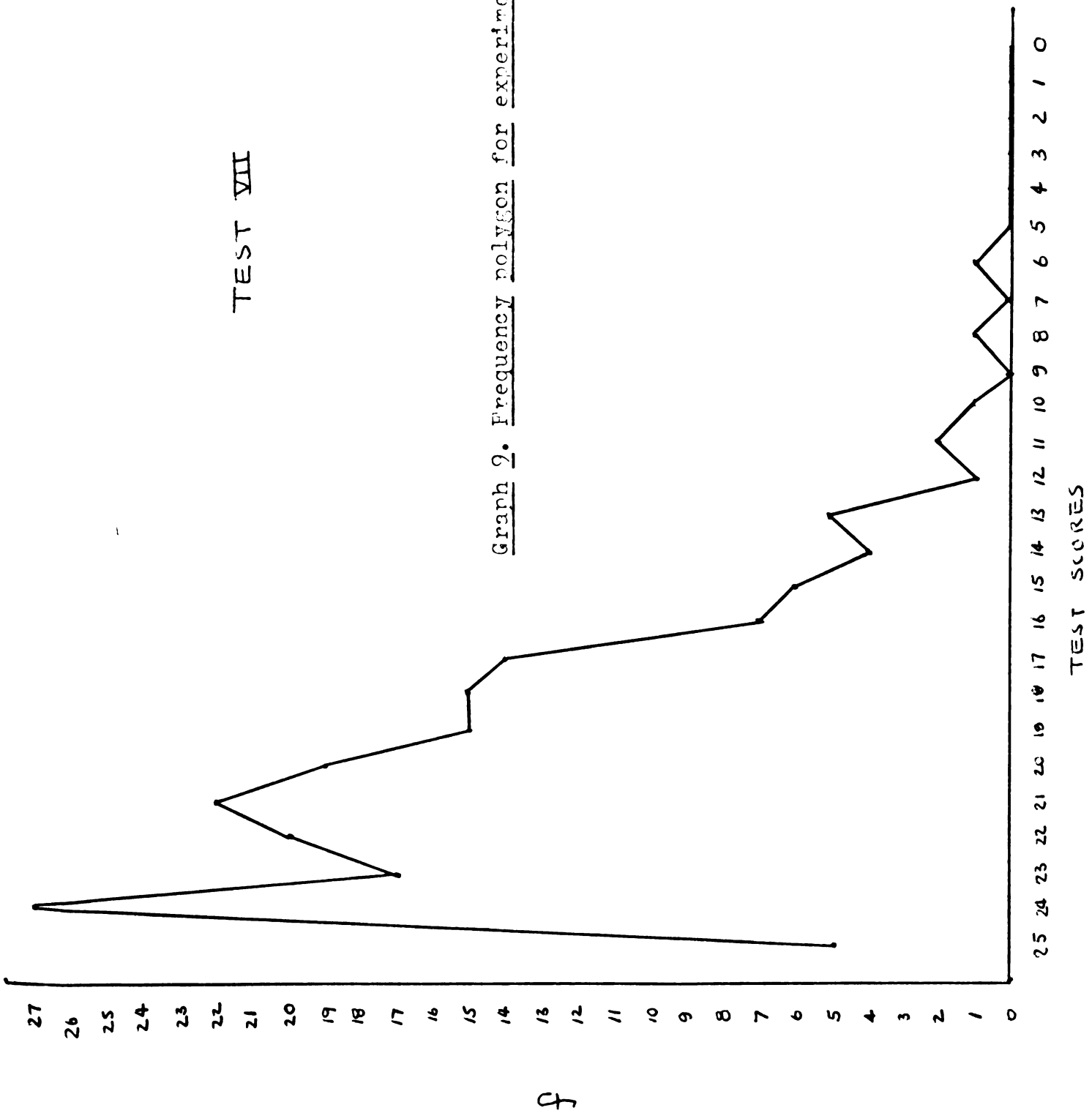
Graph 6. Frequency polygon for experimental test VI (1-25).

TEST VI (26-35)

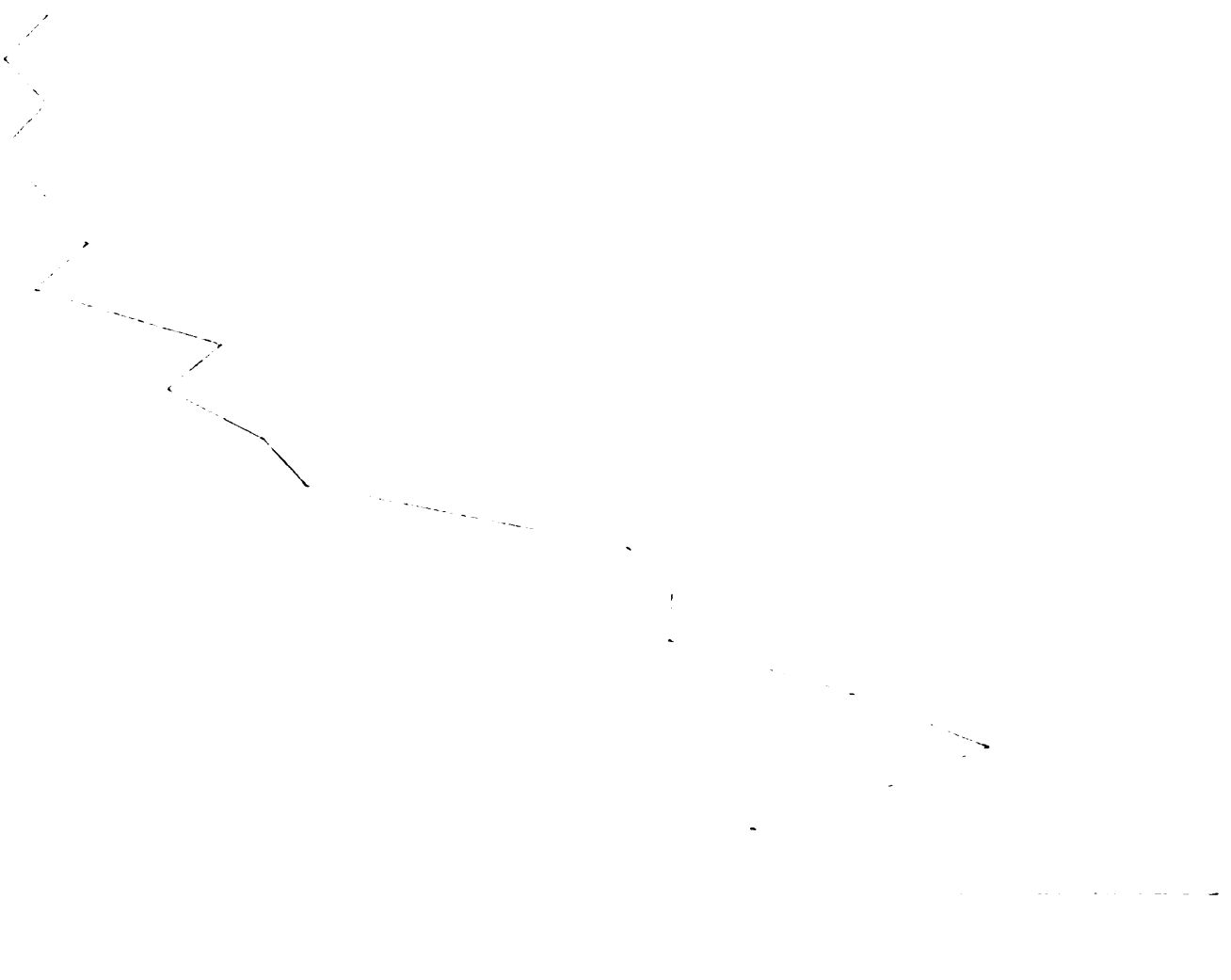


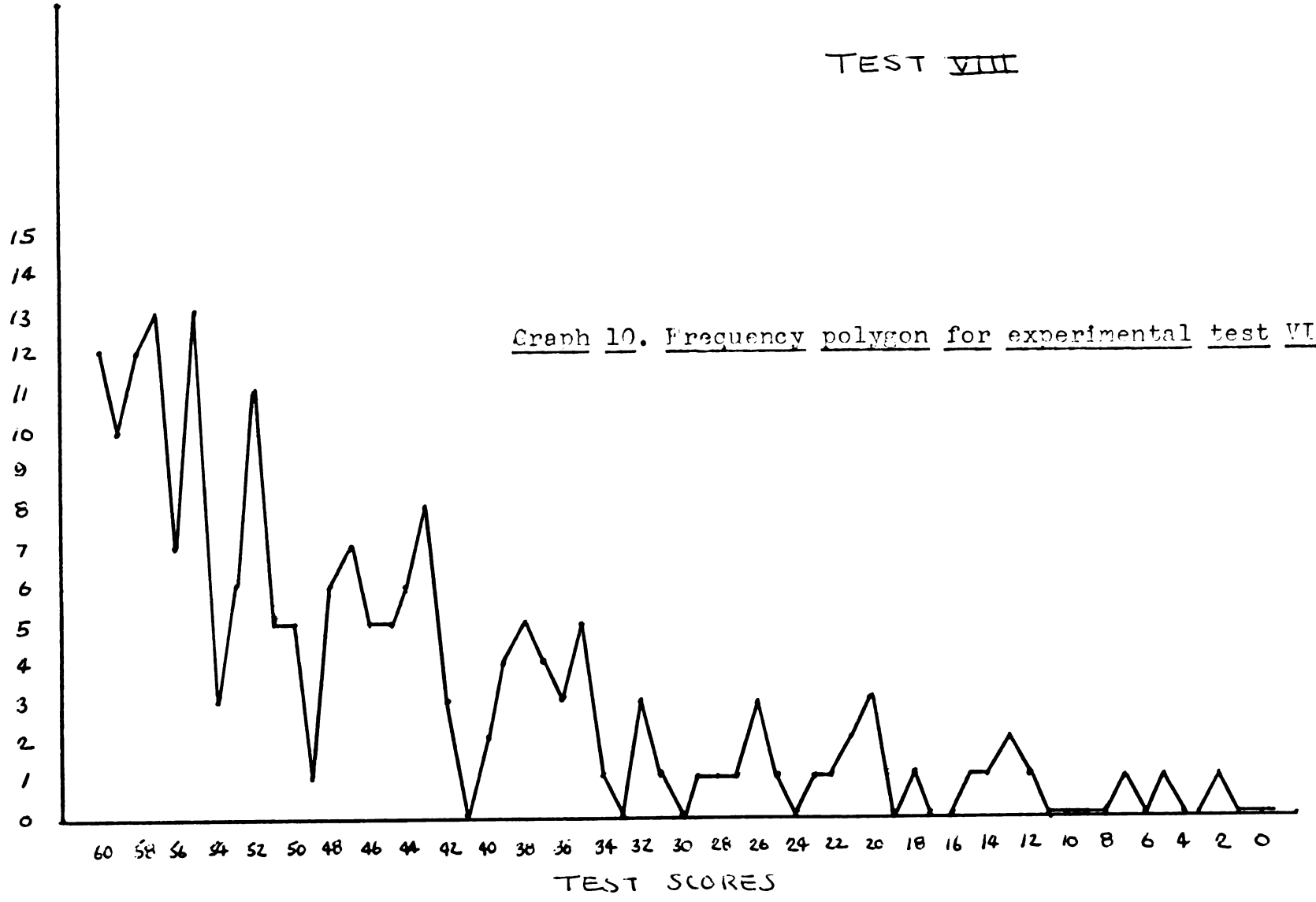
Graph 7. Frequency polygon for experimental test VI (26-35).


TEST VII (ITEMS 1-35)
Graph 8. Frequency polygon for experimental test VI (1-35).
TEST SCORES
35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22



Graph 9. Frequency polygon for experimental test VII.





Graph 10. Frequency polygon for experimental test VIII.

Table 1. Academic majors of 200 subjects

Accounting	39	Zoology	2
Marketing	34	Political Science	2
Hotel, Restaurant	28	Floriculture	1
General Business	18	English	1
Packaging	13	Chem. Engineering	1
Retailing	10	Pre-Dental	1
Business Education	9	Social Work	1
Social Science	4	Psychology	1
Interior Design	4	Economics	1
Exec. Secretarial	4	Pre-Law	1
Agriculture-Business	2	Advertising	1
Journalism	2	Dairy Production	1
Radio-Television	2	Food Science	1

Table 2. Mean, standard deviation, number of observations, and possible range for all variables.

	Variable	Mean	S. D.	Observations	Possible Range
1.	High School Size	4.810	3.393	200	0-9
2.	Number of books read per year	7.503	8.929	193	0-98
3.	Number of magazines read regularly	3.042	1.541	192	0-9
4.	Number of newspapers read regularly	1.773	0.746	194	0-9
5.	Class	3.378	0.581	196	1-6
6.	Grade Point Av.	2.427	0.391	190	0.00-4.00
7.	Reading Rate A	18.73	3.828	185	1-24
8.	Reading Rate B	13.53	4.109	185	2-30
9.	Reading Rate A+B	32.254	7.288	185	2-54
10.	Raw Comprehension	23.795	3.858	185	0-35
11.	Raw Directed Reading	9.903	3.285	185	0-20
12.	Raw Word Meaning	52.40	9.217	185	0-70
13.	Raw Sentence Meaning	40.524	6.000	185	0-50
14.	Raw Paragraph Comprehension	29.589	4.202	185	0-36
15.	Percentage of Comprehension	74.011	9.071	185	0-100
16.	Percentage of Directed Reading	92.335	11.359	185	0-100
17.	Percentage of Word Meaning	88.454	5.615	185	0-100
18.	Percentage of Sentence Meaning	90.086	7.532	185	0-100
19.	Percentage of Paragraph Comprehension	90.081	5.979	185	0-100

Table 2. (Continued).

	Variable	Mean	S. D.	Observations	Possible Range
20.	MSU Mathematics	15.600	5.622	120	0-30
21.	MSU Arithmetic	33.968	3.685	62	0-40
22.	MSU English	22.038	4.794	133	0-38
23.	MSU Reading	28.175	5.926	194	0-42
24.	CQT-Verbal	45.826	11.453	195	0-75
25.	CQT-Information	49.559	7.379	195	0-75
26.	CQT-Numerical	34.01	9.05	195	0-50
27.	CQT-Total	129.472	19.757	195	0-200
28.	Experimental Test I	17.636	5.636	195	0-25
29.	Experimental Test II	38.703	1.593	195	0-40
30.	Experimental Test III	18.646	1.356	195	0-20
31.	Experimental Test IV	19.510	4.332	194	0-25
32.	Experimental Test V	105.912	23.585	194	0-121
33.	Experimental Test VI (1-25)	17.765	4.597	183	0-25
34.	Experimental Test VI (26-35)	8.786	1.520	182	0-10
35.	Experimental Test VI (total)	26.522	5.540	182	0-35
36.	Experimental Test VII	19.825	3.673	183	0-25
37.	Experimental Test VIII	46.122	12.886	189	0-60
38.	Experimental Tests I-VIII	292.891	40.967	174	0-351

Table 3. Standard error of measurement of the test scores for experimental tests I-VIII.

Test Number	Standard Error of Measurement
I	2.28
II	1.12
III	1.12
IV	2.07
V	3.66
VI (1-25)	2.25
VI (26-35)	1.03
VI (Total)	2.53
VII	2.02
VIII	3.16

The following formula was used in computing the above standard errors of measurement:

$$S.E._{meas.} = S.D._{scores}\sqrt{1 - r_{11}}$$

where r_{11} is the reliability of the test. (27:11-20) (100:132-134)

Table 4. Reliability coefficients for eight experimental tests*

Test	Reliability Coefficient	Number of Items in Test
1. Test I	.836	25
2. Test II	.506	40
3. Test III	.314	20
4. Test IV	.772	25
5. Test V	.976	121
6. Test VI (1-25)	.757	25
7. Test VI (26-35)	.538	10
8. Test VI (1-35)	.791	35
9. Test VII	.696	25
10. Test VIII	.936	60
11. Tests I through VIII	.711	351

* These coefficients were calculated using the Kuder-Richardson Formula 21.

$$\text{rel.} = \frac{ns^2 - M(n - M)}{ns^2}$$

where n = number of items, s = standard deviation, and M = mean.

Table 5. Experimental Test I. Item correlations with the total test score.

Item Number	Correlation with total score	
1	.1655	
2	.3944	
3	.3178	
4	.3652	
5	.3552	
6	.5006	
7	.4905	
8	.2311	
9	.2528	
10	.4475	
11	.4071	
12	.5015	
13	.4404	
14	.3716	
15	.2876	
16	.3593	
17	.3359	
18	.3268	
19	.4281	
20	.3969	average item correlation with total test score is .3598
21	.4255	
22	.3003	
23	.4152	
24	.2248	
25	.2563	

Table 6. Experimental Test IV. Item correlations with the total test score.

Item Number	Correlation with total score	
1	.3369	
2	.4035	
3	.5463	
4	.5010	
5	.1353	
6	.4612	
7	.4269	
8	.4070	
9	.4489	
10	.5103	
11	.4997	
12	.5521	
13	.5594	
14	.5112	
15	.4112	
16	.4509	
17	.4901	
18	.5930	
19	.5052	
20	.5425	average item correlation with total test score is .4661
21	.4695	
22	.5393	
23	.4007	
24	.5189	
25	.4305	

Table 7. Experimental Test VI. Item correlations with the total test score.

Item Number	Correlation with total score
1	.3334
2	.3180
3	.2939
4	.5613
5	.4755
6	.5963
7	.4838
8	.5272
9	.5389
10	.4538
11	.4316
12	.3294
13	.2356
14	.2929
15	.5788
16	.3077
17	.5732
18	.5205
19	.3636
20	.3135
21	.3642
22	.4675
23	.3988
24	.4367
25	.3508
26	.3916
27	.3812
28	.3602
29	.3635
30	.2528
31	.4058
32	.4115
33	.2825
34	.4732
35	.3681

average item correlation with total test score is .4068

Table 8. Experimental Test VII. Item correlations with the total test score.

Item Number	Correlation with total score	
1	.2914	
2	.4566	
3	.4002	
4	.3508	
5	.3976	
6	.5157	
7	.5602	
8	.4485	
9	.5482	
10	.2693	
11	.2947	
12	.4837	
13	.4890	
14	.3559	
15	.3005	
16	.4403	
17	.3445	
18	.3583	
19	.3496	
20	.3673	average item correlation with total test score is .4008
21	.4156	
22	.4484	
23	.2211	
24	.5090	
25	.4027	

Table 9. Experimental Test VIII. Item correlation with total test score.

Item	r with total score	Item	r with total score
1	.2872	31	.7409
2	.2074	32	.6076
3	.3262	33	.7464
4	.3830	34	.6886
5	.5478	35	.6981
6	.5730	36	.6068
7	.6028	37	.6912
8	.5468	38	.6818
9	.3231	39	.7027
10	.4174	40	.5141
11	.5501	41	.6689
12	.5386	42	.5016
13	.5476	43	.4908
14	.5589	44	.4591
15	.4962	45	.4559
16	.5299	46	.5628
17	.3331	47	.5162
18	.3326	48	.5469
19	.4033	49	.4720
20	.5549	50	.5843
21	.6273	51	.5361
22	.6785	52	.6391
23	.6634	53	.5640
24	.6754	54	.5600
25	.6081	55	.5098
26	.4965	56	.5317
27	.3949	57	.6037
28	.4015	58	.5032
29	.6973	59	.5159
30	.7053	60	.4973

average item correlation with the total test score is .5387

Table 10. Item statistics for experimental Test I. pH = percentage of subjects scoring right in the high 27% of the total. pL = percentage of subjects scoring right in the low 27% of the total. p = number of subjects scoring right on the item in the total group. d = delta, an item difficulty index. As item difficulty increases, d increases. (see Fan, p. 3)

Item	pH	pL	p	d
1	73.58%	49.06%	56.41%	11.8
2	100.00	77.36	92.31	7.7
3	100.00	90.57	96.92	6.4
4	100.00	75.47	89.23	7.9
5	88.68	50.94	65.64	10.7
6	90.57	28.30	66.15	11.8
7	100.00	71.70	91.28	8.1
8	90.57	71.70	86.67	9.3
9	94.34	77.36	89.23	8.6
10	100.00	79.25	92.82	7.5
11	100.00	81.13	93.85	7.3
12	98.11	54.72	85.13	9.5
13	100.00	77.36	92.82	7.7
14	100.00	73.58	90.26	7.9
15	94.34	71.70	78.46	9.0
16	100.00	73.58	89.23	7.9
17	98.11	84.91	92.31	7.2
18	98.11	75.47	89.23	8.1
19	100.00	66.04	87.69	8.5
20	88.68	54.72	78.46	10.5
21	100.00	81.13	94.36	7.3

Item	pH	pL	p	d
22	90.57%	60.38%	80.00%	10.0
23	98.11	81.13	91.79	7.6
24	88.68	60.38	71.28	10.2
25	98.11	66.04	83.59	8.8

195 subjects. 53 subjects in top 27% and 53 subjects in low 27%

High and low groups are actually 27.18% of the total.

Table 11. Item statistics for Experimental Test IV.

Item	pH	pL	p	d
1	94.23%	63.46%	83.51%	9.5
2	100.00	76.92	89.69	7.7
3	100.00	50.00	77.84	9.5
4	82.69	9.62	44.85	13.5
5	42.31	26.92	32.99	14.6
6	100.00	80.77	94.33	7.3
7	98.08	73.08	90.21	8.3
8	88.46	38.46	60.82	11.5
9	96.15	48.08	72.16	10.2
10	100.00	67.31	87.11	8.5
11	84.62	13.46	50.52	13.2
12	100.00	53.85	83.51	9.3
13	100.00	57.69	85.57	9.1
14	100.00	69.23	90.21	8.2
15	88.46	28.85	56.19	12.0
16	100.00	82.69	94.33	7.1
17	100.00	44.23	74.23	9.9
18	98.08	40.38	78.35	10.3
19	100.00	69.23	90.72	8.3
20	98.08	40.38	73.71	10.3
21	100.00	76.92	93.30	7.7
22	100.00	55.77	83.51	9.2
23	98.08	69.23	89.18	8.6
24	100.00	55.77	81.96	9.2
25	100.00	76.92	92.27	7.7

194 subjects. 52 subjects in high group, 52 subjects in low group. High and low groups are each actually 26.80% of the total.

Table 12. Item statistics for Experimental Test VI.

Item	pH	pL	p	d
1	71.43%	30.61%	53.55%	12.9
2	89.80	51.02	76.50	10.6
3	91.84	67.35	84.70	9.4
4	87.76	20.41	60.66	12.5
5	83.67	24.49	53.55	12.5
6	95.92	24.49	67.21	11.5
7	97.96	57.14	83.06	9.3
8	91.84	24.49	57.92	11.9
9	100.00	59.18	87.43	9.0
10	93.88	40.82	67.21	10.8
11	71.43	20.41	43.72	13.5
12	81.63	38.78	52.46	11.8
13	89.80	61.22	76.50	10.1
14	77.55	42.86	56.28	11.9
15	85.71	8.16	45.36	13.5
16	97.96	65.31	79.23	8.9
17	95.92	24.49	63.39	11.5
18	95.92	28.57	61.20	11.2
19	95.92	63.27	81.97	9.3
20	95.92	79.59	91.26	8.1
21	95.92	71.43	87.43	8.8
22	100.00	71.43	89.07	8.2
23	97.96	55.10	79.23	9.5
24	100.00	77.55	93.44	7.6
25	95.92	63.27	83.06	9.3

Item	pH	pL	p	d
26	100.00%	93.88%	98.36%	6.3
27	100.00	87.76	96.17	6.3
28	100.00	93.88	98.36	6.3
29	95.92	67.35	80.87	9.1
30	100.00	79.59	87.43	7.4
31	87.76	32.65	57.92	11.7
32	95.92	55.10	77.60	9.8
33	100.00	89.80	95.63	6.3
34	100.00	81.63	93.99	7.2
35	100.00	93.88	98.36	6.3

183 subjects. 49 subjects in high 27% and 49 subjects in low 27%. High and low groups are actually each 26.78% of the total group.

Table 13. Item statistics for Experimental Test VII.

Item	pH	pL	p	d
1	100.00%	92.00%	96.72%	6.3
2	68.00	10.00	45.36	14.4
3	96.00	68.00	84.70	9.0
4	100.00	86.00	92.90	6.8
5	98.00	72.00	87.98	8.4
6	100.00	60.00	84.70	8.9
7	100.00	56.00	83.06	9.2
8	98.00	54.00	81.97	9.3
9	98.00	48.00	86.61	9.9
10	100.00	88.00	94.54	6.5
11	100.00	88.00	95.63	6.5
12	100.00	74.00	91.80	7.9
13	100.00	60.00	87.98	8.9
14	92.00	58.00	79.23	10.1
15	86.00	46.00	65.03	11.2
16	100.00	76.00	91.80	7.8
17	96.00	72.00	87.43	8.7
18	84.00	28.00	45.90	12.3
19	82.00	30.00	46.45	12.3
20	90.00	34.00	54.64	11.5
21	98.00	60.00	79.78	9.2
22	98.00	58.00	80.87	9.3
23	100.00	86.00	93.44	6.8
24	100.00	52.00	76.50	9.4
25	88.00	46.00	69.40	11.1

183 subjects. 50 subjects in the high 27% and 50 subjects in the low 27%. High and low groups are actually each 27.32% of the total group.

Table 14. Item statistics for Experimental Test VIII.

Item	pH	pL	p	d	
1	96.08%	74.51%	87.30%	8.5	
2	100.00	94.12	98.85	6.3	
3	90.20	50.98	66.14	10.6	Sentence 1
4	100.00	80.39	91.01	7.4	
5	96.08	74.51	87.30	8.5	
6	100.00	58.82	84.66	9.0	
7	100.00	56.86	85.71	9.1	
8	96.08	56.86	83.07	9.7	Sentence 2
9	70.59	25.49	50.26	13.2	
10	92.16	50.98	70.90	10.4	
11	100.00	27.45	59.26	10.8	
12	100.00	31.37	61.38	10.6	
13	100.00	33.33	63.49	10.5	
14	98.04	21.57	55.56	11.1	Sentence 3
15	82.35	17.65	47.62	13.0	
16	84.31	7.84	39.15	13.6	
17	100.00	92.16	97.88	6.3	
18	94.12	47.06	66.67	10.5	
19	96.08	54.90	76.72	9.8	
20	100.00	56.86	86.24	9.1	Sentence 4
21	98.04	43.14	82.01	10.1	
22	100.00	35.29	78.31	10.4	

Table 14. (continued)

Item	pH	pL	p	d	
23	100.00%	35.29%	77.25%	10.4	
24	100.00	29.41	75.13	10.7	
25	100.00	45.10	82.01	9.8	
26	100.00	60.78	88.36	8.9	Sentence 4 (continued)
27	98.04	64.71	86.24	8.9	
28	100.00	72.55	90.48	8.0	
29	100.00	54.90	87.83	9.2	
30	100.00	49.02	86.24	9.6	
31	100.00	45.10	84.13	9.8	
32	98.04	41.18	76.72	10.3	
33	100.00	37.25	81.48	10.2	
34	100.00	49.02	85.71	9.6	
35	100.00	43.14	83.07	9.9	Sentence 5
36	100.00	62.75	89.42	8.7	
37	100.00	49.02	85.19	9.6	
38	100.00	58.82	88.89	9.0	
39	98.04	49.02	85.19	9.8	
40	90.20	31.37	64.02	11.7	
41	100.00	49.02	84.13	9.6	
42	100.00	68.63	89.95	8.3	
43	92.16	29.41	58.20	11.6	Sentence 6
44	96.08	56.86	82.54	9.7	
45	98.04	37.25	65.61	10.5	
46	100.00	52.94	84.66	9.4	

Table 14. (continued)

Item	pH	pL	p	d	
47	98.04%	39.22%	68.78%	10.4	
48	98.04	37.25	70.90	10.5	
49	98.04	47.06	73.02	9.9	Sentence 6 (continued)
50	98.04	27.45	64.02	11.0	
51	90.20	25.49	56.61	12.0	
52	100.00	41.18	78.84	10.0	
53	100.00	43.14	74.60	9.9	
54	98.04	41.18	73.02	10.3	
55	92.16	43.14	75.13	10.9	
56	94.12	43.14	75.13	10.7	Sentence 7
57	100.00	41.18	76.72	10.0	
58	96.08	56.86	82.01	9.7	
59	100.00	66.67	89.42	8.5	
60	100.00	66.67	88.89	8.5	

189 subjects. 51 subjects in the high 27% and 51 subjects in the low 27%. High and low groups are actually each 26.98% of the total group.

Table 15. Test VIII. Average item statistics by sentence.

Sentence No.	Av. pH	Av. pL	Av. p	Av. d	Words
1	96.57%	75.00%	85.82%	8.2	4
2	92.48	53.92	76.98	10.0	6
3	94.12	23.20	54.41	11.6	6
4	98.86	53.08	82.27	9.41	12
5	98.94	47.66	83.23	9.77	13
6	96.86	42.16	71.43	10.3	10
7	97.82	49.24	79.31	9.8	9

Table 16. Correlation of eight experimental tests with twenty variables

	I	II	III	IV	V	VI (1-25)	VI (26-35)	VII	VIII
Gr. Pt. Av.	.24	-.09	.11	.26	.08	.18	.10	.22	.09
Read Rate A	.01	.04	-.02	.07	.05	.11	.02	.12	.10
Read Rate B	-.06	.09	.00	-.04	.03	.02	.00	.05	.05
Raw Comp.	.04	-.05	.21	.06	.17	-.05	-.08	.04	.04
Raw Dir. Rdg.	.05	.08	.10	.04	.17	.05	.14	.09	.12
Raw Wd. Mng.	.18	.12	.11	.29	.27	.23	.14	.22	.27
Raw Sent. Mng.	.23	.11	.17	.21	.28	.21	.08	.17	.20
Raw Par. Comp.	.20	.16	.29	.13	.34	.21	.13	.13	.19
% Comp.	.09	.03	.20	.02	.22	.06	-.06	.01	.06
% Dir. Rdg.	.12	-.07	.02	-.05	.07	.06	.08	-.02	.00
% Wd. Mng.	.26	.14	.20	.10	.24	.14	.15	.19	.23
% Sent. Mng.	.18	.18	.14	.21	.28	.25	.03	.19	.21
% Par. Comp.	.13	.07	.18	.05	.23	.17	.10	.08	.09
MSU Math.	.03	.07	.17	.16	.08	.18	.10	.15	.12
MSU Arith.	.11	.05	.16	-.04	.24	.14	.06	.03	.19
MSU Eng.	.14	.11	.05	.41	.32	.35	.21	.36	.26
MSU Reading	.10	.07	.26	.12	.28	.07	.02	.18	.16
CQT-Verbal	.13	.13	.19	.18	.31	.10	.10	.20	.25
CQT-Info.	.09	-.02	.13	.08	.17	.09	.06	.09	.07
CQT-Numer.	.06	.03	.17	.18	.14	.25	.14	.14	.15

Table 17. Intercorrelations of the eight experimental tests.

	I	II	III	IV	V	VI(1-25)	VI(26-35)	VII	VIII
Test I	---								
Test II	.04	---							
Test III	.19	.26	---						
Test IV	.17	-.04	.03	---					
Test V	.15	.20	.28	.23	---				
Test VI (1-25)	.34	.11	.10	.50	.34	---			
Test VI (26-35)	.36	-.02	.01	.32	.21	.52	---		
Test VII	.29	.09	.09	.60	.27	.48	.34	---	
Test VIII	.18	.21	.24	.24	.49	.36	.14	.28	---

Table 18. Rotated factor loadings for 29 major variables.

Variable	1	2	3	4	5	6	h^2
1. Grade Pt. Av.	-.09	.41	-.36	-.01	-.11	.09	.33
2. Read Rate A	.79	.06	-.12	.09	-.10	-.05	.66
3. Read Rate B	.82	-.07	.03	.00	-.03	-.07	.69
4. Raw Comp.	.25	-.08	-.76	.02	-.07	.08	.66
5. Raw Dir. Rdg.	.64	.01	-.10	.01	.12	.37	.58
6. Raw Wd. Mng.	.62	.23	-.26	.01	.33	.07	.62
7. Raw Sent. Mng.	.51	.14	-.21	-.09	.46	.19	.58
8. Raw Par. Comp.	.43	.12	-.54	-.04	.27	.25	.63
9. % Comp.	-.01	-.07	-.74	.10	-.03	.14	.58
10. % Dir. Rdg.	.20	-.04	-.04	-.01	.02	.62	.43
11. % Wd. Mng.	.04	.09	-.24	.14	.45	.22	.34
12. % Sent. Mng.	.04	.16	.00	-.15	.62	.11	.44
13. % Par. Comp.	-.12	.10	-.41	-.19	.26	.36	.42
14. MSU Math.	-.03	.20	-.11	.66	.05	-.20	.53
15. MSU Arith.	-.02	-.10	.00	.55	.23	.36	.49
16. MSU English	.03	.55	-.09	-.01	.29	-.21	.44
17. MSU Reading	.14	.14	-.58	.06	.31	-.18	.50

Table 18. (continued)

Variable	1	2	3	4	5	6	h^2
18. CQT-Verbal	.29	.19	-.48	-.10	.45	-.17	.59
19. CQT-Info.	.16	.06	-.52	.36	.12	-.14	.47
20. CQT-Numerical	.07	.15	-.05	.84	.02	.04	.74
21. Exper. Test I	-.09	.41	-.12	.00	.15	.46	.43
22. Exper. Test II	.07	-.11	.12	.07	.60	-.12	.41
23. Ex. Test III	-.10	-.06	-.30	.25	.45	.10	.38
24. Ex. Test IV	.07	.79	-.07	.03	.05	-.13	.66
25. Ex. Test V	.04	.25	-.20	.13	.57	.11	.46
26. Ex. Test VI(1-25)	.06	.71	.07	.20	.20	.22	.65
27. Ex. Test VI(26-35)	.05	.61	.11	.12	-.02	.36	.53
28. Ex. Test VII	.09	.74	-.03	.05	.15	-.05	.59
29. Ex. Test VIII	.08	.28	.02	.18	.55	.01	.42
Proportions of variance	.10	.11	.10	.06	.10	.06	

The sum is .53

Table 19. Correlations between Index 1, Index 2, and other measured variables.

Variable	Index 1	Sig.	Index 2	Sig.
Grade Pt. Av.	.22	.01	.21	.01
Reading Rate A	.01	NS	.08	NS
Reading Rate B	-.04	NS	.01	NS
Reading Rate A+B	-.01	NS	.05	NS
Raw Comprehension	.03	NS	-.05	NS
Raw Directed Reading	.03	NS	.05	NS
Raw Word Meaning	.13	.09	.20	.01
Raw Sentence Meaning	.21	.01	.18	.02
Raw Paragraph Comp.	.22	.01	.16	.04
% Comprehension	.07	NS	.06	NS
% Directed Reading	.10	NS	.07	NS
% Word Meaning	.24	.01	.16	.04
% Sentence Meaning	.15	.05	.24	.01
% Paragraph Comp.	.15	.04	.13	.09
MSU Mathematics	.17	.08	.21	.04
MSU Arithmetic	.31	.02	.33	.02
MSU English	.20	.03	.45	.01
MSU Reading	.11	NS	.05	NS
CQT-Verbal	.08	NS	.07	NS
CQT-Information	.14	.07	.03	NS
CQT-Numerical	.09	NS	.20	.01
CQT-Total	.14	.06	.15	.06
Test I	.85	.01	.36	.01
Test II	.10	NS	.06	NS
Test III	.17	.02	.08	NS
Test IV	.12	.10	.50	.01
Test V	.20	.01	.32	.01
Test VI (1-25)	.26	.01	.93	.01
Test VI (26-35)	.27	.01	.56	.01
Test VI (total)	.29	.01	.93	.01
Test VII	.28	.01	.39	.01
Test VIII	.20	.01	.32	.01
Test I-VIII (total)	.39	.01	.53	.01
Index 1			.18	.02
Index 2	.18	.02		

Significance levels are reported up to .10. Any higher than this are reported as NS.

Index 1 consists of items 1, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 16, 17, 18, 20, 21, and 23 of Test I. Index 2 consists of Items 14, 15, and 19 of Test I, item 14 of Test IV; and items 3, 4, 5, 6, 8, 9, 11, 12, 14, 15, 16, 17, 18, 19, 22, 24, 31, and 34 of Test VI. There were 17 items in Index 1 and 22 items in Index 2.

Table 20. Rotated factor loadings for the items of Tests I, IV, VI, and VII.

Test I Items	1	2	3	4	5	6	7	8	h^2
1	.08	.18	.05	.13	-.01	-.09	.11	-.13	.10
2	.02	.28	-.01	-.16	.05	.07	-.27	.33	.29
3	.00	.45	-.07	-.02	-.11	-.06	.10	.01	.23
4	-.03	.38	.00	-.04	-.03	-.03	-.02	.18	.18
5	.17	.20	-.15	.03	-.01	-.04	.02	.07	.10
6	.30	.27	.04	-.33	-.14	.00	-.05	.34	.42
7	-.05	.52	.02	-.03	-.10	-.09	-.01	-.10	.30
8	.00	.30	-.19	.00	.23	.04	-.13	-.03	.20
9	-.05	.24	.05	-.02	-.02	-.08	.10	.13	.10
10	.25	.39	-.29	-.11	.12	-.12	.02	.03	.34
11	.05	.56	-.04	.05	.06	.17	.00	-.07	.36
12	.07	.42	.06	-.23	.13	-.09	.05	.09	.27
13	.10	.40	.03	-.07	.23	.02	-.37	.10	.38
14	.08	.08	.01	-.44	.04	-.03	-.04	-.04	.21
15	-.03	.11	.01	-.27	-.04	-.05	-.10	.07	.10
16	.05	.29	-.01	-.05	.13	.04	-.01	-.09	.12
17	-.10	.33	.04	-.07	-.03	.02	.23	-.03	.18
18	.10	.32	.04	.06	.02	.00	-.03	.13	.13
19	.06	.24	.04	-.39	.07	.05	.18	.29	.33
20	.04	.32	.04	.01	.25	-.08	-.30	.14	.29
21	-.06	.55	-.07	-.03	-.04	-.14	.33	.03	.45
22	-.02	.20	.05	-.06	.01	.07	-.05	.40	.21
23	-.12	.38	.00	-.22	.16	.13	.01	-.12	.27
24	-.02	.16	.03	-.11	-.08	.02	.32	-.10	.27
25	.16	.05	-.12	-.19	-.12	.06	.27	.37	.31
Test IV Items									
1	.21	.03	.02	-.21	.12	.06	.00	-.30	.20
2	.30	.13	.13	-.04	.26	.02	.21	-.22	.29
3	.42	-.02	.14	-.13	.14	.21	.23	-.28	.41
4	.43	.06	.06	-.24	.12	.14	.05	-.14	.31
5	-.08	.09	.16	-.04	.07	-.08	.03	-.24	.11
6	.25	.22	.14	-.03	.58	.01	-.11	-.17	.52
7	.24	-.07	.04	-.03	.09	.62	-.04	-.07	.46
8	.24	-.13	.48	-.12	.14	.02	.08	.11	.36
9	.36	-.18	.23	.12	.19	.15	.18	.14	.33
10	.14	.06	.84	-.06	.14	.01	.00	-.12	.76
11	.26	.05	.06	-.10	.32	.09	.08	-.05	.21
12	.17	-.09	-.19	-.12	.58	.04	.07	-.07	.43
13	.70	-.02	-.13	-.11	.07	.02	-.06	-.02	.53
14	.22	.09	.10	-.33	.13	.23	.29	-.16	.36
15	.31	.04	-.15	-.28	.12	.10	.08	.01	.23

Table 20. Rotated factor loadings for the items of Tests I, IV, VI, and VII.

Test IV Items	1	2	3	4	5	6	7	8	h^2
16	.55	-.13	.00	-.14	.11	.08	-.01	-.01	.36
17	.35	-.12	.10	-.16	.12	.00	.28	-.20	.31
18	.25	-.18	-.16	-.24	.57	.00	.04	-.14	.53
19	.59	-.18	-.26	-.22	.10	.02	.06	-.08	.52
20	.36	-.15	.19	-.11	.34	.21	-.03	.20	.40
21	.23	.14	.09	-.06	.55	.06	-.27	.02	.47
22	.71	.02	.08	.00	.03	.05	.08	.04	.53
23	-.03	.05	.07	-.26	.38	.17	-.05	-.18	.29
24	.27	-.06	.19	.00	.43	.20	.15	.01	.36
25	.26	.00	.05	.00	.13	.54	-.09	.04	.39
Test VI Items									
1	.16	.08	-.08	-.13	.20	.03	.07	-.16	.13
2	.16	.04	.08	-.09	.20	-.18	.24	-.13	.19
3	-.01	.10	.01	-.28	.14	-.09	.01	.19	.15
4	.19	.21	.04	-.46	.20	-.03	.04	-.02	.34
5	.12	-.06	.01	-.44	.14	-.03	.22	.12	.29
6	.25	.33	.05	-.44	.13	.00	.11	-.34	.51
7	.31	.04	.05	-.31	.10	-.01	.13	-.30	.32
8	.20	.06	.07	-.39	.06	-.13	.39	-.01	.53
9	.07	.31	.00	-.48	.13	.06	.19	-.12	.40
10	.33	.03	.09	-.32	.03	-.07	.28	.08	.31
11	.34	-.17	.09	-.45	-.09	.06	.02	.05	.37
12	.21	-.04	.05	-.41	-.33	.12	-.10	-.04	.36
13	.07	.03	.10	-.01	.16	-.01	.06	-.01	.36
14	.21	-.12	.01	-.38	-.26	-.01	-.19	.02	.31
15	.11	.13	.11	-.53	.25	.13	-.04	.01	.40
16	-.17	.08	.03	-.35	-.02	-.01	.04	-.02	.16
17	.09	-.07	.10	-.50	.32	.03	.01	.02	.38
18	.02	-.14	.13	-.41	.27	.08	.03	.03	.29
19	.16	-.13	.04	-.35	.06	.22	.13	-.26	.31
20	.25	.27	.11	-.01	.11	-.06	.30	-.15	.27
21	-.04	-.06	.04	-.20	.37	.11	.13	-.16	.23
22	.17	-.05	-.02	-.37	.06	.04	.23	-.20	.27
23	.10	-.04	.04	-.22	.40	.01	.08	.20	.27
24	-.09	.25	-.29	-.33	.30	.18	.25	.23	.50
25	-.18	.14	.01	-.29	.33	.05	.22	.14	.31
26	.01	.02	.64	-.17	.14	.11	.20	.10	.52
27	.08	.04	.07	-.27	-.17	.21	.40	.18	.35
28	-.01	-.01	.85	-.17	-.10	-.01	.01	.01	.77
29	.13	.25	.09	-.10	.21	-.05	.04	-.14	.17
30	.13	.18	.03	-.09	-.03	.13	.13	.19	.13

Table 20. Rotated factor loadings for the items of Tests I, IV, VI, and VII. (continued)

Test VI Items	1	2	3	4	5	6	7	8	h^2
31	.16	-.06	-.13	-.47	.03	-.02	.11	-.17	.31
32	-.02	.19	.07	-.39	.44	.25	-.10	.11	.47
33	-.05	.05	-.33	-.09	.23	.21	.14	.09	.25
34	-.15	.01	.09	-.40	.32	.32	.22	-.05	.45
35	.18	.00	-.51	-.09	.06	.25	.08	.31	.46
Test VII Items									
1	-.07	.38	.12	.00	.38	.25	.06	-.06	.38
2	.45	.11	.09	-.27	-.02	.23	.05	-.11	.36
3	.24	.10	-.02	-.10	.19	.12	.15	-.32	.26
4	.07	-.11	.07	-.07	.07	.67	.15	-.03	.50
5	.30	.09	-.11	-.16	.18	.13	.31	.10	.29
6	.54	.09	.14	-.02	.07	.01	.19	.09	.37
7	.43	.10	-.14	.20	.44	.05	.08	-.06	.47
8	.50	.10	.02	-.16	.08	-.07	.15	-.15	.34
9	.64	.07	.11	.02	.16	.08	.00	-.01	.47
10	-.09	-.02	-.27	.00	.08	.76	-.01	-.10	.68
11	.04	.09	.08	.01	.22	.12	.20	-.23	.17
12	.28	.32	.04	-.01	.37	-.06	.00	-.21	.37
13	.68	.27	.11	.17	-.08	.14	-.01	.20	.64
14	.17	.26	.11	.19	.03	.40	-.04	.05	.30
15	.16	.14	.09	.22	.08	-.15	.31	.01	.23
16	.15	-.07	.08	-.03	.03	.74	.13	.07	.60
17	.09	.15	.82	-.09	.12	-.01	-.14	-.08	.74
18	.01	-.04	-.06	.15	.45	-.21	.23	.48	.55
19	-.03	-.03	-.08	.19	.46	-.14	.19	.43	.50
20	.05	-.07	-.07	-.11	.23	.07	.51	.01	.35
21	.30	.08	-.20	-.20	.37	-.02	-.04	-.01	.31
22	.47	.17	-.20	-.21	.09	.21	-.09	.04	.40
23	-.05	-.09	-.25	-.04	-.01	.68	.11	.04	.55
24	.20	-.02	-.10	-.01	.14	.14	.55	.00	.38
25	.12	.15	-.04	.03	.05	.04	.50	.07	.30

Table 21. Rotated factor loadings for Experimental Tests I through VIII.

Test	1	2	3	h^2
I	.09	.15	-.85	.76
II	-.06	.66	.01	.44
III	-.12	.71	-.26	.58
IV	.83	-.06	-.07	.70
V	.44	.62	.03	.58
VI (1-25)	.68	.16	-.41	.66
VI (26-35)	.44	-.08	-.64	.61
VII	.76	.08	-.22	.62
VIII	.46	.61	.06	.59
Proportions of variance	.26	.20	.16	

Table 22. Number of items in Tests I, IV, VI, and VII having their highest loadings on Factors 1-8. (from Table 20)

	Factors							
Tests	1	2	3	4	5	6	7	8
I	0	17	0	3	0	0	1	4
IV	11	0	2	1	7	2	0	2
VI (1-25)	2	0	0	16	5	0	2	0
VI (26-35)	0	1	4	2	1	0	1	1
VII	6	0	1	0	5	5	5	3

Project No. 172
Phase No. 01
Test No. _____

Name _____ Student No. _____
Last First Middle

Date _____ Age in Months _____

Sex _____ Approximate number of students in your high school? _____
M-F

Approximately how many books do you read in a year? (not counting required texts and outside readings for courses) _____

How many magazines do you read regularly? _____

How many newspapers do you read regularly? _____

What is your class? (Fr, Soph, Jr, Sr, Grad) _____

What is your major? _____ G.P.A. (All Univ.) _____

(Please do not write below this line)

****

Age _____	CQT Verbal _____
Sex _____	CQT Info. _____
G.P.A. _____	CQT Numerical _____
High School _____	CQT Total _____
Class _____	MSU Math _____
Books _____	MSU Arith _____
Magazines _____	Test I _____
Newspapers _____	Test II _____
Reading RAte _____	Test III _____
Comprehension 1 _____	Test IV _____
Comprehension 2 _____	Test V _____
Comprehension Total _____	_____
MSU English _____	_____
MSU Reading _____	_____

(1)

Test I

Instructions for Test I.

Read the short introduction. Read each sentence that follows and judge whether or not the sentence could be an English sentence. The sentences contain nonsense words so do not attempt to find a "meaning" for the unfamiliar words. Base your judgements on whether or not the sentences have the tune, the sound, the rhythm, the pattern of English sentences.

****

Back in 1922, the late Heywood Broun, who is not remembered primarily as a boxing writer, wrote a durable account of a combat between the late Benny Leonard and the late Rocky Kansas for the lightweight championship of the world.

1. Leonard was the glatest pretitioner of the ora, Kansas less a duff, rutomistic foral. A. Eng?_____B. Not Eng?_____

2. In the roily down, Kansas and Leonard. A. Eng?_____B. Not Eng?_____

3. Broun was purloinedly refrected. A. Eng?_____B. Not Eng? _____

4. A madical in soltices, he was abjective in the pars. A. Eng?_____B. Not Eng?_____

5. Kansas lape him vink of Gertrude Stein. A. Eng?_____B. Not Eng?_____

6. The Kansas in a round. A. Eng?_____B. Not Eng?_____

7. He widled that he had been merp tethful as a bilp. A. Eng?_____B. Not Eng?_____

8. There is still a vick in stryle, and movation narries a basty callop. A. Eng?_____B. Not Eng?_____

9. I of Broun's worls in the vater Kears. A. Eng?_____B. Not Eng?_____

10. The burrent Rock is gauche and inabburate, but he is a drepfully sereal vitter. A. Eng?_____B. Not Eng?_____

11. The prelominative nazure of this assef has been well slaped by Pierce Egan. A. Eng?_____B. Not Eng?_____

12. Broun with vertimate inclations. A. Eng?_____B. Not Eng?_____

13. Egan said, "He porrelles a bequisite above all the ert that neaching can achiele for any boxer." A. Eng?_____B. Not Eng?_____

14. This is drue not only of Marciano's rigler kand but of his lesk kand, too.
A. Eng?_____B. Not Eng?_____

15. Egan doubted changing approved only.
A. Eng?_____B. Not Eng?_____

16. He would have aggoved of Marcian's fyle.
A. Eng?_____B. Not Eng?_____

17. The grampion has an abbarently unpimited golarity for parament.
A. Eng?_____B. Not Eng?_____

18. Piffing or moving to the fide, and moving yack, are innogrations of the late eigelanth densury.
A. Eng?_____B. Not Eng?_____

19. Mr. Egan these tactics in boxers of vorpal contions.
A. Eng?_____B. Not Eng?_____

20. He dord tofe into his grame of leverance.
A. Eng?_____B. Not Eng?_____

21. Archie Moore, who bilinates in Morpal, and elirates in Boggell, is a Brounian rather than an Eganite.
A. Eng?_____B. Not Eng?_____

22. Since the rine of Marciano who has been acrile.
A. Eng?_____B. Not Eng?_____

23. Moore, who has been acrile, has nuffered the dangs of a sutreme exglopent of belanto.
A. Eng?_____B. Not Eng?_____

24. I regreived signed of a note in his borunt.
A. Eng?_____B. Not Eng?_____

25. A nellow who has as much fyle as Moore in a ving by the dorpels with many ciller drollefs.
A. Eng?_____B. Not Eng?_____

Test I Total _____

Test II

INSTRUCTIONS FOR TEST II.

Choose the words in the answer list that will fit in the appropriate blanks in the sentence. Each word is used only once. There are more words in the answer list than are needed to fill the blanks. Work carefully. Write clearly. Be sure that your answers are in the proper blanks.

****

1. In the _____(1), a _____(2) _____(3)ed _____(4)ly.

   1 is _____ 3 is _____

   2 is _____ 4 is _____

   Answers: 1. of 2. be 3. summer 4. for 5. boy 6. some
   7. slow 8. walk

2. The _____(5)er _____(6) _____(7)ed a _____(8)to _____(9)

   5 is _____ 8 is _____

   6 is _____ 9 is _____

   7 is _____

   Answers: 1. be 2. me 3. hat 4. by 5. of 6. tall 7. toss
   8. lad 9. only 10. within

3. The _____(10)est _____(11) _____(12)s _____(13)ish.

   10. is _____ 12. _____

   11. is _____ 13. _____

   Answers: 1. with 2. act 3. fool 4. girl 5. were 6. tall
   7. very 8. should

4. _____(14)were _____(15) a _____(16)by the _____(17).

   14. is _____ 16. _____

   15. is _____ 17. _____

   Answers: 1. I 2. committee 3. talked 4. given 5. very 6. we
   7. here 8. prize

(4)

5. The _____ s were _____ by a _____ _____.
18 19 20 21

18 is _____ 20. is _____

19 is _____ 21 is _____

Answers: 1. some 2. wealthy 3. of 4. given 5. man 6. be
7. because 8. book

6. The _____ _____ s _____ _____.
22 23 24 25

22 is _____ 24 is _____

23 is _____ 25 is _____

Answers: 1. are 2. of 3. new 4. some 5. should 6. table
7. that 8. here

7. The _____ s _____ ed _____ _____.
26 27 28 29

26 is _____ 28 is _____

27 is _____ 29 is _____

Answers: 1. chairman 2. with 3. him 4. since 5. elect 6. taller
7. voter 8. very

8. _____ of _____ _____ ed that _____ _____.
30 31 32 33 34

30 is _____ 32 is _____

31 is _____ 33 is _____

34 is _____

Answers: 1. man 2. to 3. b 4. consider 5. us 6. were
7. foolish 8. all

9. The _____ _____ s _____.
35 36 37

35 is _____ 36 is _____ 37 is _____

Answers: 1. very 2. young 3. bite 4. of 5. are 6. dogs
7. for 8. much

10. The _____ _____ _____.
38 39 40

38 is _____ 39 is _____ 40 is _____

Answers: 1. hit 2. were 3. are 4. ball 5. are 6. Fred 7. by
8. any

(5)

Test II Total _____

TEST III

(6)

INSTRUCTIONS FOR TEST III.

Choose the answers from the answer list that stand for the symbols used in the sentences. A word from the answer list will stand for each symbol used. There are more answer words than are needed.

****

A. He is a dredfully severe hitter ▽ either hand. The predominative nature □ this asset has been stated △ Egan and Mallory □ the old London prize ring.

1. ▽ stands for _____

2. □ stands for _____

3. △ stands for _____

Answers: 1. with 2. very 3. of 4. by 5. in 6. the 7. any 8. for

B. I did not think □ he could bring it off, △ I wanted to be there ▽ he tried.

4. □ stands for _____

5. △ stands for _____

6. ▽ stands for _____

Answers: 1. when 2. by 3. to 4. but 5. up 6. around 7. dawn 8. that

C. When I heard △ the boys had been made □ the fight, at the Yankee Stadium, I shortened my stay abroad ▽ order not to miss the encounter ⧖ the two heroes.

7. △ stands for _____

8. □ stands for _____

9. ▽ stands for _____

10. ⧖ stands for _____

Answers: 1. for 2. so 3. of 4. to 5. because 6. that 7. by 8. in

D. □ London △ the night ▽ September 13th, a week before the date set ⧖ for the encounter, I tried to get my eye □ ⧖ fight-watching ✳ attending a bout ## the White City greyhound track.

11. □ stands for _____

12. △ stands for _____

13. ▽ stands for _____

14. ⧖ stands for _____

15. ✳ stands for _____

16. ## stands for _____

Answers: 1. on 2. for 3. in 4. by 5. to 6. of 7. down 8. at

E. △ I had engagements □ kept me in England ▽ a few days ✳ the Encounter, I had no opportunity to visit the training camps.

17. △ stands for _____

18. □ stands for _____

19. ▽ stands for _____

20. ✳ stands for _____

Answers: 1. that 2. before 3. until 4. who 5. by 6. some 7. by 8. because

Test IT total _____

TEST IV

Instructions for Test IV

This is a "part-of-speech" test. Choose the part of speech of each underlined word from the answer list and put the number of the answer on the appropriate blank. Be sure that the numbers match. That is, be sure that the number of the underlined word that you are identifying matches the blank in which you write your answer number.

****

A. I (1) reflected (2) with (3) satisfaction (4) that (5) old (6) Moore (7) could (8) have (9) whipped (10) all (11) four (12) principals (13) on (14) that (15) card (16) within (17) fifteen (18) rounds (19).

1. is a _____
2. is a _____
3. is a _____
4. is a _____
5. is a _____
6. is a _____
7. is a _____
8. is a _____
9. is a _____
10. is a _____
11. is a _____
12. is a _____
13. is a _____
14. is a _____
15. is a _____
16. is a _____
17. is a _____
18. is a _____
19. is a _____

B. Dempsey may have been (20) a great (21) champion (22), but (23) he had (24) less to beat than Marciano (25).

20. is a _____
21. is a _____
22. is a _____
23. is a _____
24. is a _____
25. is a _____

Answers: 1. noun 2. prounoun 3. verb 4. adverb 5. adjective 6. conjunction 7. preposition.

Test IV Total _____

Test V

Instructions for Test V

This is a translation test. Read the introductory sentence. Then read sentence one. Sentence one (and the following sentences) are real sentences made up of real words that sound something like the words they are to be translated into. There are exactly enough blanks to use in the translation. Here is an example

Marry hatter ladle limb, itch fleas worse widest snore.

Marry	Hatter	ladle	limb,	itch	fleas
Mary	had	a little	lamb,	its	fleece
1.	2	3 4	5	6	7

worse	widest	snore.
was	white	as snow
8	9	10 11

Please work carefully and write clearly.

****

Once upon a time there was a little girl who lived with her mother in a little cottage on the edge of a large dark forest.

1. Disc ladle gull orphan worry putty rat cluck wetter ladle rat hut.

Disc	ladle	gull	orphan	worry	putty
1	2	3	4	5 6	7

rat	cluck	wetter	ladle	rat	hut.
8	9	10	11 12	13	14 .

2. Fur disc raisin pimple orphan colder Ladle Rat Rotten Hut.

Fur	disc	raisin	pimple	orphan	colder
15	16	17	18	19	20 21

Ladle	Rat	Rotten	Hut.
22	23	24	25 .

3. Wan moaning Ladle Rat Rotten Hut's murder colder inset:

Wan	moaning	Ladle	Rat	Rotten	Hut's
26	27	28	29	30	31

murder	colder	inset:
32	33	34 35 36

4. "Ladle Rat Rotten Hut, heresy ladle basking winsome burden barter and shirker cockles."

Ladle	Rat	Rotten	Hut,	heresy		
37	38	39	40	41	42	43

ladle	basking	winsome	burden	barter
44	45	46	47 48 49	50

an	shirker	cockles.
51	52	53

5. "Tick disc ladle basking tutor cordage offer groin-murder hoe lifts honor udder site offer florist."

"Tick	disc	ladle	basking	tutor	
54	55	56	57	58	59

cordage	offer	groin-murder	hoe	lifts
60	61 62	63	64	65

honor	udder	site	offer	florist."
66 67	68	69	70 71	72

6. "Dun stopper laundry wrote!"

"Dun	stopper	laundry	wrote!
73	74	75	76 77

7. "Dun stopper peck floors!"

Dun	stopper	peck	floors.
78	79	80 81	82

8. "Dun daily-doily inner florist an dun stopper torque wet strainers!"

Dun	daily-doily	inner	florist	an	dun
83	84	85 86	87	88	89

stopper	torque	wet	strainers.
90 91	92	93	94

9. "Hoe-cake, murder," resplendent Ladle Rat Rotten Hut an stuttered oft.

"hoe-cake,	murder,"	resplendent	Ladle	Rat
95	96	97	98	99

Rotten	Hut	an	stuttered	oft.
100	101	102	103	104

10. Honor wrote tutor cordage offer groin-murder, Ladle Rat Rotten Hut mitten anomalous woof.

Honor	wrote	tutor	cordage	offer
105 106	107	108 109	110	111 112

groin-murder,	Ladle	Rat	Rotten	Hut
113	114	115	116	117

mitten	anomalous	woof.
118 119	120	121

Test V Total _____

Test VI

Name _____ Student No. _____ Sec. ____ Date_____

I. A. In the frammis a morgrant arablint daskaped.

B. The borpal toggen profrumes the glasplart.
1 2 3 4 5 6

1. What word in sentence B has the same function as morgrant in sentence A? _____

2. What word in sentence B has the same function as daskaped in sentence A? _____

II. A. The crellest frop delanders a bront into the horent.

B. In the plastof, a nufrant relograzed a bilant sindel.
1 2 3 4 5 6 7 8 9

3. What word in sentence B has the same function as delanders in sentence A?_____

4. What word in sentence B has the same function as crellest in sentence A? _____

5. What word in sentence B has the same function as frop in sentence A? _____

III. A. The lorfest nuffer of the blint will cranzale the wiltrof.

B. The crawfletch vorpaled the parler meff in the frammis.
1 2 3 4 5 6 7 8 9

6. What word in sentence B has the same function as lorfest in sentence A? _____

7. What word in sentence B has the same function as blint in sentence A? _____

8. What word in sentence B has the same function as wiltroff in sentence A? _____

9. What word in sentence B has the same function as cranzale in sentence A? _____

10. What word in sentence B has the same function as nuffer in sentence A? _____

IV. A. While craffis bolomered the snaffle, crenshaw drozened by the morpaler falet.

B. Plorent clafels the nark of an effel, so zarkis glins the
1 2 3 4 5 6 7 8 9 10 11
glonest whiffle.
12 13

11. What word in sentence B has the same function as falet in sentence B? _____

12. What word in sentence B has the same function as Craffis in sentence A? _____

13. What word in sentence B has the same function as morpaler in sentence A? _____

14. What word in sentence B has the same function as bolomered in sentence A? _____

V. A. The poller vink dapazes spoothly.

B. Soon a narlest harpen glomed.
1 2 3 4 5

15. What word in sentence B has the same function as spoothly in sentence A? _____

16. What word in sentence B has the same function as poller in sentence A? _____

17. What word in sentence B has the same function as dapazes in sentence A? _____

18. What word in sentence B has the same function as vink in sentence A? _____

VI. A. What morvent was fradled by the dorl?

B. Which fram of the dornels grabels the blint?
1 2 3 4 5 6 7 8

19. What word in sentence B has the same function as dorl in sentence A? _____

20. What word in sentence B has the same function as fradeled in sentence A? _____

21. What word in sentence B has the same function as morvent in sentence A? _____

22. What word in sentence B has the same function as by in sentence A? _____

VII. A. Will zepran degradle the mopril ?

B. In the orgrant, bramel frandors a flant.
1 2 3 4 5 6 7

23. What word in sentence B has the same function as mopril in sentence A? _____

24. What word in sentence B has the same function as degradle in sentence A? _____

25. What word in sentence B has the same function as zepran in sentence A? _____

VIII.A. In the afternoon a beautiful rainbow appeared.

B. The tall boy chases the bus.
1 2 3 4 5 6

26. What word in sentence B has the same function as <u>beautiful</u> in sentence A? _____

27. What word in sentence B has the same function as <u>appeared</u> in sentence A? _____

IX. A. The youngest player tossed a ball through the window.

B. On the stage, a workman paints an old table.
1 2 3 4 5 6 7 8 9

28. What word in sentence B has the same function as <u>tossed</u> in sentence A? _____

29. What word in sentence B has the same function as <u>youngest</u> in sentence A? _____

30. What word in sentence B has the same function as <u>player</u> in sentence A? _____

31. What word in sentence B has the same function as <u>window</u> in sentence A? _____

X. A. The old horse walks slowly.

B. Soon a bright light appeared.
1 2 3 4 5

32. What word in sentence B has the same function as <u>slowly</u> in sentence A? _____

33. What word in sentence B has the same function as <u>old</u> in sentence A? _____

34. What word in sentence B has the same function as <u>walks</u> in sentence A? _____

35. What word in sentence B has the same function as <u>horse</u> in sentence A? _____

TEST VII

Name _____ Student # _____ Sec # _____ Test # _____

A. He bolomered (1) with rezation (2) that siller (3) Bronk (4) could have dalazed (5) all borunts (6) on that cravel (7) within vorent (8) rowls (9).

1 is a _____ 6 is a _____

2 is a _____ 7 is a _____

3 is a _____ 8 is a _____

4 is a _____ 9 is a _____

5 is a _____

B. Derzey (10) may have calovered (11) a marler (12) charlion (13), but glorer (14) had less to vint (15) than Marlinio (16).

10 is a _____ 13 is a _____ 16 is a _____

11 is a _____ 14 is a _____

12 is a _____ 15 is a _____

C. The burrent (17) Charlie is gauche (18) and inabburate (19), but he is a drepfully (20) sereal (21) vitter (22).

17 is a _____ 19 is a _____ 21 is a _____

18 is a _____ 20 is a _____ 22 is a _____

D. Glovan (23) was purloinedly (24) refrected (25).

23 is a _____ 24 is a _____ 25 is a _____

ANSWERS: 1. noun 2. pronoun 3. verb 4. adverb 5. adjective 6. conjunction 7. preposition.

Please put the number of the correct answer in the appropriate blank.

TEST VIII

NAME _____ STUDENT # _____ SEC. ___ TEST # _____

1. Debt's jest hormone nurture

Debt's	jest	hormone	nurture
1	2	3	4

2. Oil ketchup wetter letter.

Oil	ketchup		wetter		letter.
5	6	7	8	9	10

3. A nervous sausage bag ice!

A	nervous	sausage		bag	ice!
11	12	13	14	15	16

4. Wile your wrestling, yore kin mocker bets an washer dashes.

Wile	your	wrestling,	yore	kin
17	18	19	20	21

mocker	bets	an	washer	dashes.
22	23	24	25	26	27	28

5. Yore kin leader ~~horace~~ hearse toe warder, butcher cannon maggot drank.

Yore	kin	leader	hearse	toe	warder,
29	30	3k	32	33	34	35

butcher	cannon	maggot	drank.
36	37	38	39	40	41

6. Jest snuff doze ~~[illegible]~~ orders combing firmer putty rat rat roaches.

Jest	snuff	doze	orders	combing	firmer
42	43	44	45	46	47

putty	rat	roaches.
48	49	50	51

7. Heresy rheumatic starry offer former's dodder.

Heresy			rheumatic	starry
52	53	54	55	56
offer		former's	dodder.	
57	58	59	60	.

Perceived Structure of Written Utterances
Project 172-Phase 01
Department of Communication
Michigan State University
Director: Mr. Hepler

CODE SHEET

Card 01

Column	Item	Code
1-3	Project Number	172
4-5	Phase Number	01
6-11	Student Number	000000-999999
12	Test Number 1	1
13-37	Sentences 1-25	0=wrong 1=right 9=omitted or data not available
38-39	Test 1 Total (scored R-W)	00-25 99=data not available
40	Test Number 4	4
41-65	Questions 1-25 of Test 4	0=wrong 1=right 9=data not available
66-67	Test 4 Total	00-25 99=data not available
79	Complete?	0=not complete 1=complete
80	Card Number	1

Card 02

Column	Item	Code
1-3	Project Number	172
4-5	Phase Number	01
6-11	Student Number	000000-999999
12	Test Number 6	6
13-47	Questions 1-35 Test 6	0=wrong 1=right 9=data not available

Card 02 (cont.)

Column	Item	Code
48-49	Test Number 6 Total	00-35 99=data not available
50	Test Number 7	7
51-75	Questions 1-25 of Test 7	0=wrong 1=right 9=data not available
76-77	Test 7 Total	00-25 99=data not available
79	Complete?	0=no 1=yes
80	Card number 2	2

Card 03

Column	Item	Code
1-3	Project Number	172
4-5	Phase Number	01
6-11	Student Number	000000-999999
12	Test Number 8	8
13-72	Word number of Test 8	0=wrong 1=right 9=data not available
73=74	Test 8 Total	00-60 99=data not available
79	Complete?	0=no 1=yes
80	Card number 3	3

Card 04

Column	Item	Code
1-3	Project Number	172
4-5	Phase Number	01
6-11	Student Number	000000-999999
12	Sex	0=female 1=male
13-15	Academic major	not used

Card 04 (continued)

Column	Item	Code
16	High School Size	0=under 200 1=201 to 400 2=401 to 600 3=601 to 800 4=801 to 1000 5=1001 to 1200 6=1201 to 1400 7=1401 to 1600 8=1601 to 1800 9=over 1800
17-18	Number of books read per year. (Not counting textbooks)	00-98 99=data not available
19	Number of magazines read regularly.	0-9
20	Number of newspapers read regularly.	0-9
21	Class	1=freshman 2=sophomore 3=junior 4=senior 5=graduate 6=other
22-23	Major	see detailed code
24-26	Grade Point Average	000-400 999=data not available
27-28	Reading Rate A	00-24 99=data not available
29-30	Reading Rate B	00-30 99=data not available
31-32	Reading Rate A+B	00-54 99=data not available
33-34	Raw Comprehension A+B	00-35 99=data not available
35-36	Raw Directed Reading	00-20 99=data not available
37-38	Raw Word Meaning	00-70 99=data not available

Card 04 (Continued)

Column	Item	Code
39-40	Raw Sentence Meaning	00-50 99=data not available
41-42	Raw Paragraph Comprehension	00-36 99=data not available
43-45	% Comprehension A+B	000-100 999=data not available
46-48	% Directed Reading	000-100 999=data not available
49-51	% Word Meaning	000-100 999=data not available
52-54	% Sentence Meaning	000-100 999=data not available
55-57	% Paragraph Comprehension	000-100 999=data not available
58-59	MSU Mathematics	00-30 99=data not available
60-61	MSU Arithmetic	00-40 99=data not available
62-63	MSU English	00-38 99=data not available
64-65	MSU Reading	00-42 99=data not available
66-67	CQT Verbal	00-75 99=data not available
68-69	CQT Information	00-75 99=data not available
70-71	CQT Numerical	00-50 99=data not available
72-74	CQT Total	000-200 999=data not available
79	Complete?	0=no 1=yes
80	Card number 4	4

Card 05

Column	Item	Code
1-3	Project Number	172
4-5	Phase Number	01
6-11	Student Number	000000-999999
12-13	Test Number 1 Total	00-25 99=data not available
14-15	Test Number 2 Total	00-40 99=data not available
16-17	Test Number 3 Total	00-20 99=data not available
18-19	Test Number 4 Total	00-25 99=data not available
20-22	Test Number 5 Total	000-121 999=data not available
23-24	Test Number 6 (1 thru 25 total).	00-25 99=data not available
25-26	Test Number 6 (26 thru 35 total.)	00-10 99=data not available
27-28	Test Number 6 Total	00-35 99=data not available
29-30	Test Number 7 Total	00-25 99=data not available
31-32	Test Number 8 Total	00-60 99=data not available
33-35	Tests 1 thru 8 Total	000-351 999=data not available
79	Complete?	0=no 1=yes
80	Card Number 5	5

* *

Notes:

1. Deck 1 is original data with omitted and unavailable data coded as 9, 99, or 999.

2. Deck 2 has the missing data recoded to the mean category for the distribution.

MICHIGAN STATE UNIVERSITY LIBRARIES
3 1293 03085 1707