DOCUMENT RESUME

ED 060 757 FL 003 045

AUTHOR Oller, John W., Jr.
TITLE Assessing Competence in ESL: Reading.
PUB DATE 1 Mar 72
NOTE 25p.; Revised version of a speech presented at the Sixth Annual TESOL Convention, Washington, D.C., March 1, 1972

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS College Students; Correlation; *English (Second Language); Eye Movements; Foreign Students; Language Instruction; Language Proficiency; Language Research; *Language Skills; *Language Tests; Linguistic Competence; Reading Comprehension; *Reading Skills; Reading Speed; *Reading Tests; Second Language Learning; Test Construction; Test Reliability

ABSTRACT

Results from research with eye movement photography (EMP) are discussed with a view to defining differences between native-speaker and non-native reading processes. The greatest contrast is in terms of the duration of eye fixations; non-native speakers at the college level require about as much time for a fixation as an average native-speaker at the third grade level. Various tests of reading skill are discussed and correlations with other tests are given. The hypothesis is advanced that high correlations between tests of listening, speaking, reading, and writing are an indication of test validity. Support for tests which can easily be constructed by classroom teachers is provided. (Author/VM)

ASSESSING COMPETENCE IN ESL: READING[1]

John W. Oller, Jr.

Southwestern Cooperative Educational Laboratory, Inc.
Albuquerque, New Mexico

Abstract

Results from research with Eye Movement Photography (EMP) are discussed with a view to defining differences between native speaker and non-native reading processes. The greatest contrast is in terms of the duration of eye fixations; non-native speakers at the college level require about as much time for a fixation as an average native speaker at the third grade level. Various tests of reading skill are discussed and correlations with other tests are given. The hypothesis is advanced that high correlations between tests of listening, speaking, reading, and writing are an indication of test validity. Support for tests which can easily be constructed by classroom teachers is provided.

---

[1]This paper was presented in a much abbreviated form at the TESOL Convention (Washington, D. C.; March, 1972) at the invitation of David Harris, Chairman of the section entitled "Measuring Proficiency in Reading: From Theory to Practice."

I. Why Measure Reading Skill?

Of all the skills required by students of English as a second language surely none is more important to success in college-level course work than the ability to read at a reasonable rate and with comprehension. Unfortunately, most ESL courses at the college level that I have observed, in my opinion, either ignore the reading skill, or deal with it inadequately. The objective of the ESL course (or sequence of courses) is usually to "enable the student to achieve his full potential in college-level course work," or some similar safely vague goal. Strangely enough, however, the student is rarely given practice in reading, or if he is given a reading assignment, the material to be covered is usually far beneath the level of practically any college text. Foreign students are frequently expected to learn to read college-level prose by practicing material appropriate in complexity for fourth or fifth graders and in content suitable only for idiots.

The reason for mentioning the foregoing is that tests of reading and other ESL skills are sometimes used with little or no reference to their institutional purpose. The first question to be considered is: what do you intend to do about it once you find out whether a student can or cannot read college-level material? Or, stated differently, how will the more proficient reader be treated differently than the less proficient one? Obviously, I cannot answer these questions for all institutions. In fact, I cannot even answer them for UCLA. I call them to your attention, however, because they are of great importance to a total instructional program in ESL which incorporates a testing procedure for assessing reading skill. It is my opinion that reading should not only be tested, but it should also be taught.

The data that I will discuss comes primarily from research recently completed or still in progress in connection with the UCLA English as a Second Language Placement Examination (ESLPE). We will consider some data from eye movement photography (EMP) and from various tests traditionally associated with reading and some others besides. Inferences will be made particularly concerning the relationship between aspects of reading and other language skills.

II. Comparing Natives and Non-Natives Using EMP

Although some people have argued that reading is only incidentally a visual process [redacted], it is nevertheless a form of visual information processing, and one way of examining it is to measure various dynamics of eye movements during the reading process. A sophisticated photographic technique for doing this known as eye movement photography (EMP) has been available since the 1950's (cf. Taylor, 1960, and his references). Research with this technique shows clearly that, contrary to some popular views, the eye cannot sweep across a page in a smooth and rhythmic flow. Reading actually takes place by a series of jerky movements. The speed at which the eye moves from one fixation to the next is such that virtually all of the information processing occurs during the fixations rather than during transitions between them. Among the parameters of eye movements which have been investigated (Taylor, [redacted] 1960) are average fixations per hundred words, regressions per hundred words (i.e., the average number of times the student looked back to re-read), span of recognition (i.e., the average number of words taken in per fixation), average duration of fixation, average number of words per minute (with at least 70% comprehension as determined by a multiple choice test on the material read). An important question to be considered in a discussion

of the reading proficiency of non-native speakers of English is just how their skill compares to that of native speakers. EMP is certainly one way of getting some data on this question.

In the summer of 1970, Tullius and I did a study with EMP employing fifty foreign students at the University of California, Santa Barbara. We compared these fifty non-natives against the norms for some 12,000 native speakers on the various EMP measurements. The results of this comparison are given in Figure 1. For the fifty non-natives tested, there was a remarkable contrast with the norms for native speakers. While many of the

- - - - - - - - - - - - - - -

Insert Figure 1 about here

- - - - - - - - - - - - - - -

non-natives nearly equaled college-level native performance in terms of fixations, word span, and words per minute; on the average, their performance in terms of duration of eye fixations was about third- or fourth-grade level. There were few individual exceptions to this. Also, in number of regressions, it is interesting that non-natives did not deviate in performance significantly from college-level native speakers. From all of this, it is possible to conclude with some assurance that the real contrast between native and non-native performance is in the speed with which they process verbal information in short-term memory. Even when the group of non-natives was partitioned into several sub-groups of differing levels of proficiency (cf. Oller and Tullius, 1972), the non-natives (even those who had studied at an institution where English was the medium of instruction) differed significantly from natives in terms of duration of fixation ($p < .05$ by Duncan's New Multiple Range Test; Dixon, 1970).

III. Correlation: Part A. Its Interpretation

One of the possible myths of theories of second language learning is that the processes of speaking and understanding are essentially distinct from reading and writing. This assumption has led experts in language testing to pre-suppose that devices for the assessment of the supposedly separate skills of speaking, understanding, reading, and writing should be different in type and should produce substantially different results. This has in its turn led to the conclusion that if the parts of a proficiency examination such as the Test of English as a Foreign Language (TOEFL, produced by Educational Testing Service, Princeton, New Jersey), for example, are very highly intercorrelated the validity of the test is thereby reduced. Today, I would like to advance the alternative hypothesis that low correlations between tests intended to measure the various language skills are, in fact, indicative of a lack of validity in the tests themselves and, conversely, high correlations are indicators of basic test strength.

In order to clarify the last point for teachers who may not be familiar with the technical meaning of "correlation," I will try to give an adequate, though technically incomplete, layman's description of what the statistic of (Pearson's product-moment) correlation means. Perhaps the most important characteristic of any test is the extent to which it spreads the examinees on a scale. In interpreting this scale, the score of any given student can be said to be high or low only in relation to some other hypothetical or real score used as a point of reference. Usually this is the arithmetic average, otherwise known as the mean. The amount of deviation from the mean for any given score can be algebraically represented as a square. By adding up all of the squares of the deviations from the mean for a set of scores, and dividing by the number (n) of students

(sometimes, n-1), we derive what statisticians refer to as the variance of those scores. If the same individual takes two or more tests, it is possible to correlate the results of the tests by determining to what extent they tend to produce overlapping variances. In other words, we ask the question, to what extent do the scores on tests A and B, A and C, B and C, and so on, tend to vary in the same direction from their respective means. The amount of overlap between any pair of tests taken by the same subjects can be expressed as a correlation in the form of a percentage.

The interpretation of correlations is probably best explained by example. For instance, if we find that test A correlates at the 90% level with test B, this means that 81% of the variance contained in test A is also contained in (or is predicted by) test B. We derive the 81% by squaring 90%. This correlation and variance overlap is represented in Figure 2. On the other hand, if we find a correlation of 30%, this means that only about 9% of the variance contained in test A is also present in test B. The latter correlation and variance overlap is represented in Figure 3. Correlation expressed as a percentage, then, is roughly parallel

- - - - - - - - - - - - - - - - - -

Insert Figures 2 and 3 about here

- - - - - - - - - - - - - - - - - -

to the concept of deviation from the mean while the square of the correlation is parallel to the notion of variance. We must remember that a high correlation between two tests does not necessarily indicate that the two tests are measuring the same thing. It merely arouses this suspicion. However, it does indicate, in general, the extent to which one test may be used as a predictor of scores on the other(s) when similar populations are tested.

Part B. Correlating EMP Measurements with the UCLA ESLPE

In addition to the comparison between natives and non-natives, described above in connection with the EMP study, Tullius and I also correlated the EMP measurements with the results on the UCLA ESLPE Form 1 which is used by UCSB. This examination consisted of five parts: vocabulary (synonym matching in the context of a sentence), grammar, sound perception (minimal pairs in ambiguous sentence contexts), dictation, composition, and oral interview. The correlations between each of these subtests and the various measurements derived from the EMP test are given in Table I.

- - - - - - - - - - - - - -

Insert Table I about here

- - - - - - - - - - - - - -

Negative correlations in the table indicate that when one score tended to be high, the other tended to be low. For example, the greater the number of fixations, the lower the score on the ESLPE. Hence, we get a negative correlation (- 55%) between fixations and the test total for the ESLPE. It is interesting that the correlations between the dictation, the interview, and the various EMP measures are generally higher than with any of the other parts of the ESLPE, and also, the correlation between the interview and the dictation is nearly 90%. Another interesting fact is that the highest correlation between duration and part scores on the ESLPE are with dictation and oral interview (-24% and -22% respectively). The latter fact seems to stem from the general difficulty foreign students have in rapid processing of language, visually in reading, auditorily in taking dictation, and productively in speaking. All of this data taken as a whole shows a fairly substantial predictive relationship between

measures of reading and other integrative language skills. It may be noteworthy too that the correlation between the vocabulary test and the measures of reading are generally lower than for the oral interview and the dictation (cf. Table I).

Part C. Dictation and Integrative Skills Testing

Until the fall of 1969, Form 1 of the UCLA ESLPE (in several versions which had been used up to that time) consisted of five parts: a dictation; a composition on any one of three topics for which the student was allowed approximately 45 minutes; a vocabulary test of the synonym matching variety using words in a discourse context; a sound discrimination task requiring that the student distinguish between minimal pairs in sentence contexts, and finally a grammar test in which the student selected from three sentences the one that seemed most acceptable. (This was, in fact, pretty much the same as the test used at Santa Barbara and mentioned earlier in connection with the Oller and Tullius study.) None of the subtests in itself purported to measure the reading skill as we usually think of it with the possible exception of the vocabulary section. By intercorrelating the five parts of that examination for a sample of 100 subjects who took it in the fall of 1968, I stumbled onto something that will probably surprise few classroom teachers but that had to that time been largely overlocked by professional test writers (cf. Oller, 1971a). This accidental discovery set the stage for the research relating to the measurement of reading which we will consider briefly.

The surprising result of the correlation analysis was that the dictation predicted scores on all parts of the ESLPE Form 1 better than any other single part (see Table II). I suppose that some testing experts would

- - - - - - - - - - - - - - -

Insert Table II about here

- - - - - - - - - - - - - - -

have used this result as a basis for eliminating the dictation from the test. However, I was ignorant of many of the accepted dogmas of language testing at that time so I took this result to mean that the dictation was yielding more information about language proficiency than any other part of the test. This also seemed to me to be a good argument in favor of tests which measure global language skills all in a bunch rather than attempting to segregate out the bits and pieces of behavior that are sometimes presumed to make up the totality of language skill.

Ever since this initial discovery concerning dictation, the main thrust of the UCLA ESLPE has been progressing in the direction of a more integrative test. Among the changes that we made were the following: (1) we threw out the section on sound discrimination [for several reasons that I will not mention here]; (2) we eliminated the composition because of scoring difficulties; (3) we lengthened the vocabulary test; (4) we included a sentence paraphrase task and a paragraph reading task as measures of reading skill; and (5) we lengthened the grammar test to include items that required the ordering of words, phrases, and clauses to create a meaningful sequence to fit a given context. This left us with a four-part ESLPE Form 2: vocabulary, grammar, reading, and dictation. By modifying the test in this way, we were also able to do some experimentation with other test forms. Results of intercorrelations from four different versions of the ESLPE Form 2 which were administered to several different groups of subjects are given in Tables III-VI.

- - - - - - - - - - - - - - - - -

Insert Tables III-VI about here

- - - - - - - - - - - - - - - - -

There were a few important differences in the various versions of Form 2. In Form 2A (see Table III), the dictation in question, due to administrative problems, had to be given by six different people in six different locations who no doubt used different reading rates, etc. In spite of this, there is very high intercorrelation between dictation and the other test parts. The reading score, however, is a slightly better predictor of the test total (86%) than is the dictation (84%). In practically every case the intercorrelations of parts indicated in Table III are 80% or better.

In Table IV for Form 2B, although the intercorrelations between dictation and the other test parts is not as high as it was for 2A, the dictation predicts the total score well enough that for some purposes it could have validly been substituted for the whole test. Also, from Table V, it is apparent that Form 2C performed similarly to 2B.

Test Form 2D involved some substantial changes. All of the multiple-choice items had five alternatives instead of four as in previous examinations. The grammar test was lengthened to 80 items instead of 40. Half of these items were devoted to problems of word, phrase, or clause order while the other half required that the subject select the most appropriate of five possible words, phrases, or clauses to fill in a given blank in a larger structure. The vocabulary test was changed so that students had to match a word with one of five possible synonyms, but in a sentence context (previous forms used items in isolation). This change is probably the explanation for the improved correlation with the reading score (85%) and

the higher correlation with the total (90%). Perhaps the most interesting result of the analysis displayed in Table VI is that each part of the test yielded a 90% correlation with the total score except for the dictation, and the correlation there was 88%.

All of the foregoing taken as a whole suggests that there is a substantial underlying connection between reading and the other language skills tested by the sub-parts of the UCLA ESLPE. This is somewhat surprising in that the mode of perception and/or production varies markedly on the different tasks required by these sub-tests.

Part D. Cloze Tests: Reading and Other Skills

Another of the tests traditionally recognized as a measure of reading skill that has recently received considerable attention at UCLA and elsewhere is the cloze procedure. This is a method of constructing a fill-in blank test by deleting every fifth, sixth, or seventh word from a passage of prose. The task for the student is to fill in the blanks. This test has been used with native speakers as an index of the readability of passages of prose, and also as a measure of reading comprehension. When Darnell (1968) correlated a cloze test with the total and part scores on the Test of English as a Foreign Language (TOEFL) and found the highest part score correlation with the Listening Comprehension section (73%), many people, especially the experts, were surprised (cf. Spolsky, 1970). Since a cloze test is usually considered a test of reading, one might expect it to correlate best with the reading and vocabulary sections of the TOEFL.

Darnell's result has been replicated indirectly several times in research with the UCLA ESLPE. Conrad and I found that cloze scores correlated better with dictation (82%) than with any other section of the UCLA ESLPE Form 2C (Oller and Conrad, 1971). I assume that the dictation on

the UCLA ESLPE yields information closer to the TOEFL Listening Comprehension score than any other section on the latter test. In a later study (Oller, 1971b), utilizing the ESLPE Form 2A Revised along with three cloze tests (see Table VII), again the highest correlations in every case (using the acceptable word scoring method discussed below) were between cloze scores and dictation. The next highest in each case was between cloze tests and reading.

- - - - - - - - - - - - - - -

Insert Table VII about here

- - - - - - - - - - - - - - -

Another technique which was employed in an attempt to determine the interrelation between the different test types used in the study with cloze tests was partial correlation. By this method, it is possible to statistically test the amount of overlap between two variables while controlling for a third, or possibly a third and fourth, or a third, fourth, and fifth, etc. Figure 4 roughly illustrates the meaning of the technique. The overlap between A and C, for instance, may be partialed out while checking the overlap between A and B. Or, more control variables may be used. If only one control is used, the result is referred to as a first order partial; if two, it is called a second order partial; and so on.

- - - - - - - - - - - - - - -

Insert Figure 4 about here

- - - - - - - - - - - - - - -

Table VIII gives the results of this technique when applied to the correlations of the cloze tests and part scores on the UCLA ESLPE Form 2A Revised (Oller, 1971b). In every case, the correlation between cloze scores and dictation is superior to any other part of the examination.

- - - - - - - - - - - - - - -

Insert Table VIII about here

- - - - - - - - - - - - - - -

The correlation with vocabulary partials out completely when three control variables are used.

From this same study, it was also determined that of five scoring methods investigated the most valid and economical procedure for scoring a cloze test is to count any response that conforms to all of the contextual constraints of the passage as correct, and to count incorrect any response that violates any constraint whatsoever. In a follow-up research project, Bowen, Dien, Mason and I (1972) tested speakers of Thai, Vietnamese, and English on cloze tests in each of these languages. An interesting result of the latter study was that translating a passage from one language into another makes possible the construction of cloze tests of equivalent difficulty in both languages.

A possible ramification of this result if it is in fact correct, is that it will be possible to make an approximate comparison between the performance of students in a foreign language class and native speakers of that foreign language as follows. Translate a test passage from the foreign language into English. Construct a cloze test in English by deleting, say, every seventh word. Test a group of native speakers of English of approximately the same IQ, grade in school, etc., as the students in the foreign language class. Then make a cloze test of the original passage in the foreign language. (If you deleted every seventh word in the English passage, you should also delete every seventh word in the foreign language passage.) Now, give the test to your foreign language class. The score of the native speakers of English on the English cloze test will provide

an approximate estimate of the score that native speakers of the foreign language would have made on the FL cloze test. Comparing the scores of the foreign students with those of the native speakers of English should indicate how close they are in skill to native speakers of the foreign language of comparable age, socio-economic level, etc.

IV. Testing in Relation to Teaching Reading

Now I would like to return momentarily to the issue of teaching reading skill in relation to its testing. The research that I have reported on in this paper suggests that there is very likely a much closer relation between the four skills than has been assumed traditionally. The assumption that you should test one point of language in one specific skill area at a time seems less credible than it once was. Similarly, the notion that a foreign language is best learned through practice involving a progression through a series of discrete pattern drills each offering a slightly more complex structure of the language is suspect. Should a reader be graded for difficulty to the extent that it becomes emptied of attention-holding context? Which is more important: controlling the structures taught or tested, or controlling the content presented so that the student progresses through a series of more and more complex conceptualizations? There are a number of excellent reading programs prepared for native speakers of English at the various grade levels which do not make any attempt at the rigid structural controls sometimes demanded by "experts" in applied linguistics. One of these programs that deserves the attention of teachers of ESL is the complete set of materials prepared by Science Research Associates. In my judgment, these materials are far superior to many currently being used in ESL classes. (They were called to my attention by Tamar Goldmann, one of our teaching assistants at UCLA.)

If testing should, as I have argued here, progress in the direction of instruments designed to measure more integrative skills, this has some serious implications for teaching. One way of formulating a goal is to say that the student will have mastered the following structures by such and such a date. This works fairly well at the elementary levels, but at the more advanced levels, a more reasonable goal might be to ask that the student be able to read such and such a passage with 90% comprehension as determined by some test of comprehension. If the teacher chooses to use devices such as cloze tests and dictations both as teaching and testing techniques, some important questions need to be dealt with. What is the effect of repeated practice in taking dictation (cf. Valette, 1964)? What is the effect of repeated practice in taking cloze tests? Are either or both of these really learning experiences? Is there significant transfer from the learning in one of these tasks to the other? (Harriet Kirn, one of our teaching assistants at UCLA is working on these questions.) Does improvement in the ability to do cloze tests successfully also tend to result in improved reading comprehension? Speaking fluency? Listening comprehension? Writing skill? There is a great deal of research to be done in classroom situations on questions such as these. I hope that many of you will become interested in doing it.

Figure 1

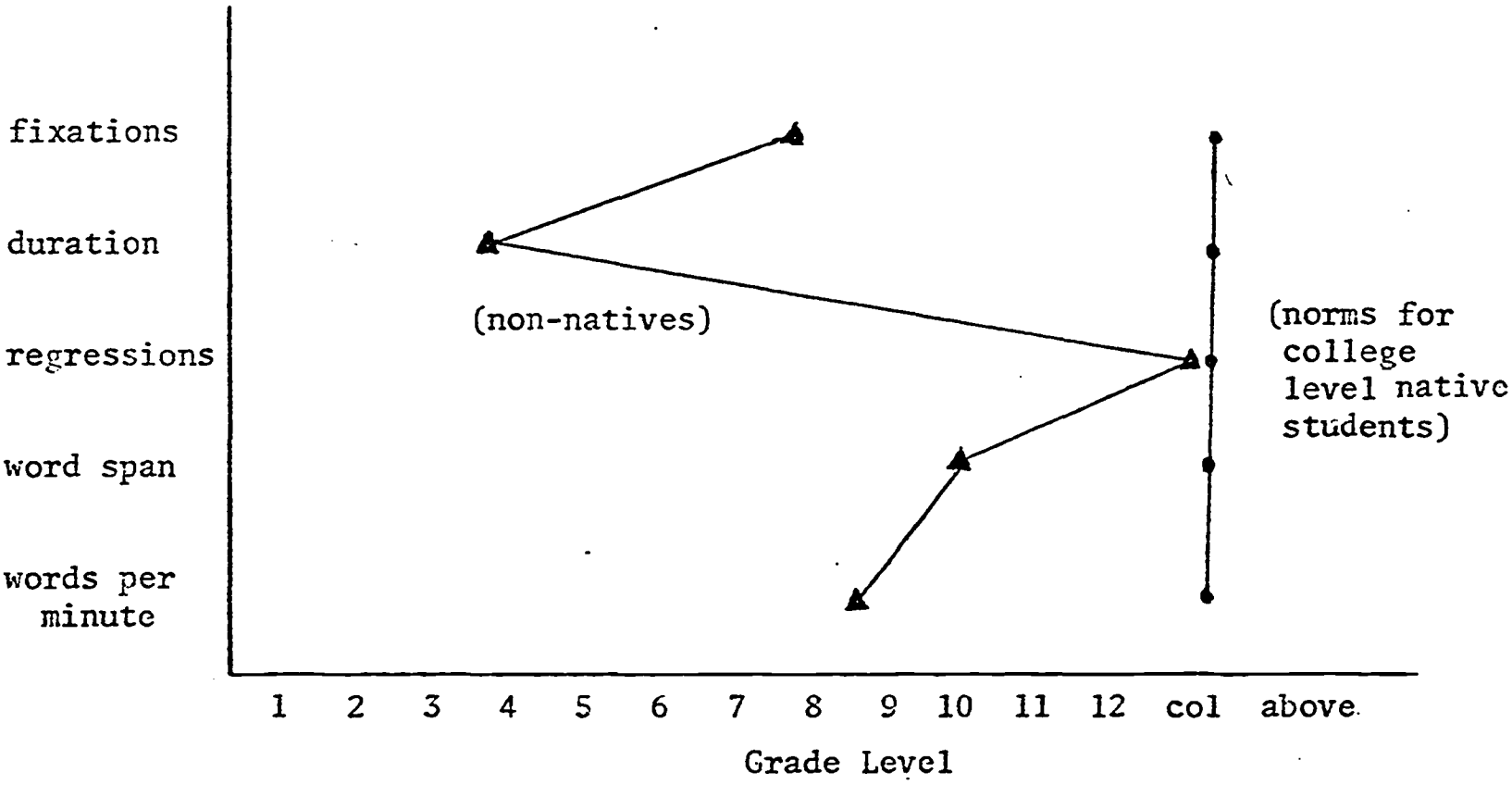

Figure 1. A comparison of average EMP measurements based on national norms for the various grade levels of natives (12,000 in all) and scores of 50 non-natives.

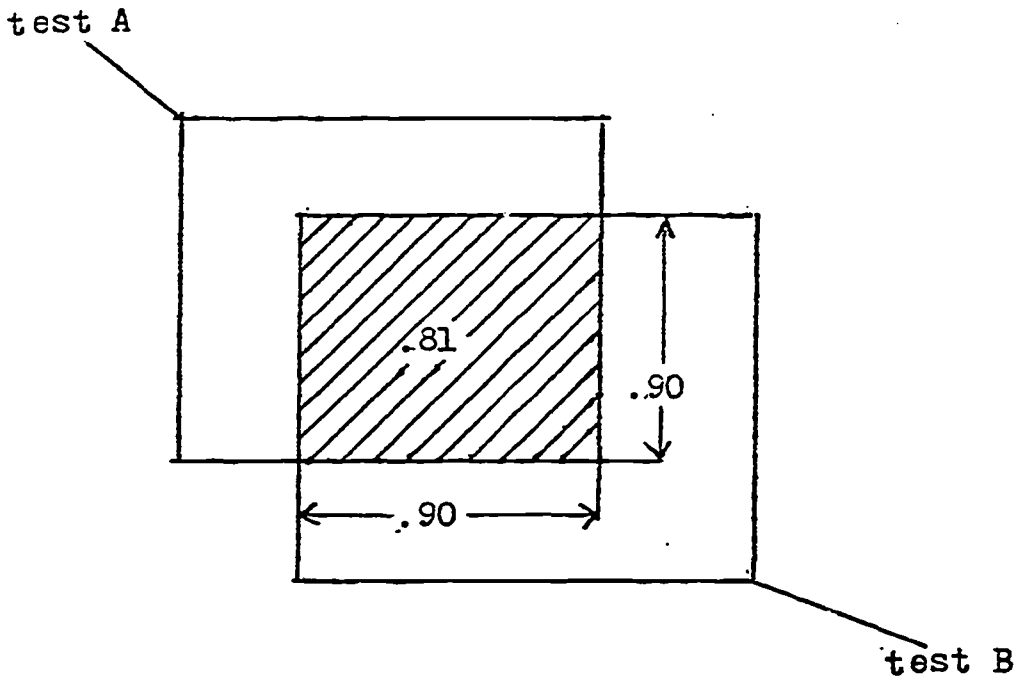

Figure 2. Correlation of .90 between tests A and B represented as a .81 ($.90^2$) overlap in test variance.

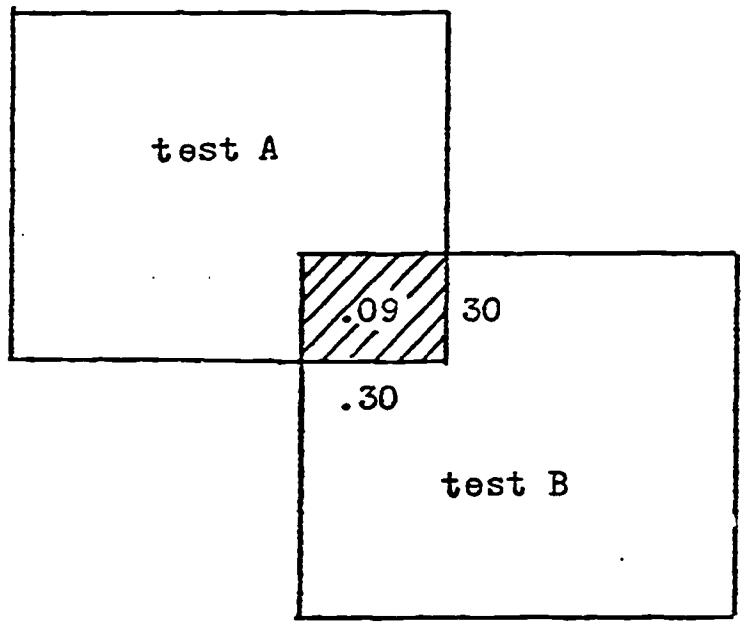

Figure 3. Correlation of .30 between tests A and B represented as a .09 ($.30^2$) overlap in test variance.

Figure 4.

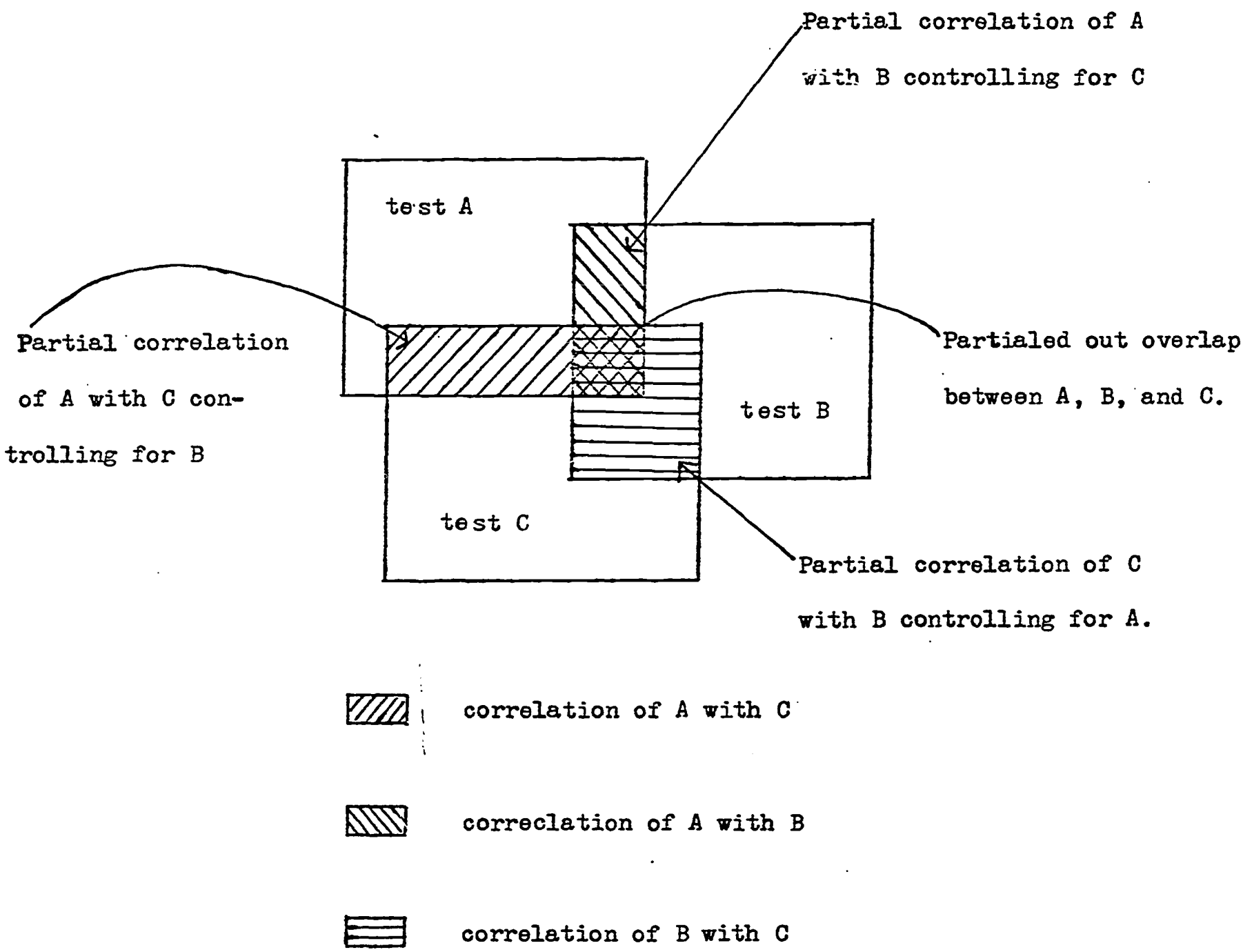

Figure 4. Partial correlation represented as the overlap in variance between tests A and B minus the overlap of test C.

	Fixations	Durations	Regressions	Word Span	Words/ Minute	Vocabulary	Grammar	Sound Perception	Dictation	Composition	Interview	Total
EMP												
1 Fixations		.12	.82	-.91	-.78	-.45	-.47	-.30	-.49	-.44	-.54	-.55
2 Duration			.10	-.17	-.50	-.08	-.02	-.01	-.24	-.07	-.22	-.14
3 Regressions				-.69	-.58	-.36	-.51	-.37	-.40	-.46	-.44	-.53
4 Word Span					.90	.50	.49	.31	.55	.49	.56	.60
5 W/Minute						.44	.41	.24	.58	.48	.57	.58
ESLPE												
6 Vocabulary							.65	.38	.66	.54	.62	.76
7 Grammar								.60	.65	.73	.64	.85
8 Sound Perception									.54	.44	.60	.62
9 Dictation										.73	.89	.90
10 Composition											.63	.88
11 Interview												.86

Table I. Correlations between all EMP measurements with each of the part scores and the total on the UCLA ESL Placement Examination, Form 1 (N=50).

Table II. Intercorrelations of the Part Scores and Total Score for the UCLA ESLPE Form 1 (N=100 subjects entering in the fall of 1968)

	Composition	Grammar	Vocabulary	Phonology	Total
Dictation	.69	.64	.67	.57	.86
Composition		.55	.51	.53	.88
Grammar			.58	.50	.78
Vocabulary				.45	.77
Phonology					.69

Table III. All Non-Redundant Intercorrelations of Part Scores and Total Score for the UCLA ESLPE Form 2A (N=291 subjects entering in the fall of 1969-70).

	Reading	Grammar	Vocabulary	Total
Dictation*	.83	.79	.80	.84
Reading		.83	.81	.86
Grammar			.83	.82
Vocabulary				.81

*Due to an administrative problem the dictation was given by six different people in six different rooms--this probably explains the fact that it does not do as well here as in later test versions.

Table IV. All Non-Redundant Intercorrelations of Part Scores and Total Score for the UCLA ESLPE Form 2B (N=227 subjects entering in the winter of 1969-70)

	Reading	Grammar	Vocabulary	Total
Dictation	.73	.73	.69	.96
Reading		.73	.80	.84
Grammar			.72	.82
Vocabulary				.81

Table V. All Non-Redundant Intercorrelations of Part Scores and Total Score for the UCLA ESLPE Form 2C (N=120 subjects tested in the spring of 1969-70)

	Reading	Grammar	Vocabulary	Total
Dictation	.81	.75	.77	.96
Reading		.80	.83	.90
Grammar			.76	.86
Vocabulary				.87

Table VI. All Non-Redundant Intercorrelations of Part Scores and Total Score for the UCLA ESLPE Form 2D (N=145 subjects entering in the spring of 1970-71)

	Reading	Grammar	Vocabulary	Total
Dictation	.70	.73	.76	.88
Reading		.78	.85	.90
Grammar			.71	.90
Vocabulary*				.90

*Changed to items in sentence contexts.

Table VII. All Non-Redundant Intercorrelations of the UCLA ESLPE 2A Revised Part Scores and Total Score with Three Cloze Tests of Different Levels of Difficulty (Easy N=132, Medium N=134, Hard N=129 foreign students entering UCLA in the fall of 1970).

Cloze I (Easy)					
	Reading	Grammar	Vocabulary	Cloze I	Total
Dictation	.71	.74	.73	.76	.96
Reading		.67	.74	.71	.83
Grammar			.71	.73	.84
Vocabulary				.63	.85
Cloze I					.80

Cloze II (Medium)					
	Reading	Grammar	Vocabulary	Cloze II	Total
Dictation	.74	.71	.74	.84	.96
Reading		.74	.80	.78	.87
Grammar			.76	.76	.84
Vocabulary				.75	.87
Cloze II					.89

Cloze III (Hard)					
	Reading	Grammar	Vocabulary	Cloze III	Total
Dictation	.75	.80	.71	.85	.97
Reading		.73	.69	.82	.84
Grammar			.74	.80	.87
Vocabulary				.71	.81
Cloze III					.89

Table VIII. First-, Second-, and Third-Order Partials for Three Cloze Tests combined correlated with the UCLA ESLPE Form 2A Revised.

UCLA ESLPE	First-Order Partials		Second-Order Partials		Third-Order Partials	
	Cloze	Control Variable	Cloze	Control Variables	Cloze	Control Variables
I. Vocabulary	.27	II	.09*	II, III		
	.25	III	.10*	II, IV	.00*	II, III, IV
	.20	IV	.05*	III, IV		
II. Grammar	.41	I	.33	I, III		
	.40	III	.23	I, IV	.19	I, III, IV
	.29	IV	.20	III, IV		
III. Reading	.42	I	.34	I, II		
	.41	II	.29	I, IV	.26	I, II, IV
	.34	IV	.28	II, IV		
IV. Dictation	.53	I	.42	I, II		
	.48	II	.45	I, III	.37	I, II, III
	.50	III	.38	II, III		

*Failed to reach significance at $p < .001$.

References

Darnell, D. The Development of an English Language Proficiency Test of Foreign Students Using a Clozeentropy Procedure. ERIC ED 024-039.

Dixon, W. J. (Ed.) Biomedical Computer Programs. Berkeley, University of California Press, 1970.

Oller, J. W. Jr. "Dictation as a Device for Testing ESL Proficiency," English Language Teaching, 25 (1971) 254-9. (a)

_____ "Scoring Methods and Difficulty Levels for Cloze Tests of ESL Proficiency," Modern Language Journal (1971, in press). (b)

_____, J. D. Bowen, T. T. Bien, and V. W. Mason. "Cloze Tests in English, Thai and Vietnamese," Language Learning, 22 (1972, in press).

_____, and C. Conrad. "The Cloze Technique and ESL Proficiency," Language Learning, 21 (1971), 183-195.

_____, and J. Tullius. "Reading Skills of Non-native Speakers of English," IRAL (1972, in press).

Spolsky, B. "Recent Research in TESOL," TESOL Quarterly, 3 (1969), 355.

Taylor, S. E. Eye Movement Photography with the Reading Eye. Huntington, New York: Educational Development Laboratory, 1960.

Valette, R. M. "The Use of Dictée in the French Language Classroom," Modern Language Journal 48 (1964), 431-4.