DOCUMENT RESUME

ED 278 700 TM 870 117

AUTHOR Kerstiens, Gene
TITLE Time-Critical Reading Comprehension Tests and Developmental Students.
PUB DATE Apr 86
NOTE 18p.; Paper presented at the Annual Meeting of the American Educational Research Association (67th, San Francisco, CA, April 16-20, 1986).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Community Colleges; *Developmental Studies Programs; Higher Education; *Reading Comprehension; Reading Difficulties; *Reading Tests; Remedial Reading; Response Style (Tests); Scores; Student Behavior; Student Placement; *Testing Problems; *Timed Tests
IDENTIFIERS *Speededness (Tests)

ABSTRACT

This study investigated the effect of timed testing on developmental students' scores and developed data that might better reveal students' reading and testing behaviors. Three tests were administered to provide six variables, including the Cooperative English Test, Form 1A (Coop 1A); Cooperative English Test, Form 1B (Coop 1B); and Degrees of Reading Power, PA-2. Subjects were 95 students in four community college developmental reading classes conducted during the fall 1984 semester. Results showed that the relationship between developmental students' scores on untimed reading tests is positive. However, the correlations between scores on a time-critical reading comprehension test and two untimed reading tests are negligible, and these low correspondences are attributable to the timing variable. Timing or the speed set it elicits not only effects scores, but also the manner in which a test measures. It was learned that developmental students' scores on a time-critical test are profoundly affected by students' speed of response. It is apparent that time-critical reading comprehensive tests, which critics view as unsuitable for student populations as a whole, are even less appropriate for developmental students whose rate behaviors are comparatively slower. (LMO)

TIME-CRITICAL READING COMPREHENSION TESTS AND DEVELOPMENTAL STUDENTS

Gene Kerstiens
El Camino College

A paper delivered at the Annual Meeting of the American Educational Research Association, San Francisco, April 16-19, 1986

BACKGROUND

Virtually all students entering college must submit to placement testing, which typically includes a standardized test designed to assess students' reading comprehension skills. These reading comprehension tests are usually timed and employ difficult passages that require exacting reading and thinking skills in order to answer the comprehension items that accompany them (Rounds 1984). Moreover, the number of passages and test items are sufficiently numerous so that test compilers' expectations are that only a gifted or randomly-marking minority will finish the test during the time allotted. On such time-critical tests, rate behaviors quite obviously influence students' scores and therefore their academic placement in, advancement through, or even preclusion from programs or courses. The consequence of these test scores can be academically profound.

PERSPECTIVES AND OBJECTIVES

The preponderance of studies performed on college/adult populations during the last thirty-five years supports the position that speed of

comprehension reading tests that employ difficult passages and require exacting reading and thinking skills lack validity (Barton 1975; Davis 1962; McDonald 1971; Preston and Botel 1951; Rankin 1962; Stetson 1982). Regardless of this evidence and virtually none to the contrary, "Ironically, the increasing popularity of certain tests seems to be inversely related to the negative comments of critics " (Tillman 1977).

Most of the studies dealing with the effect of time constraints on the reading test scores of college/adult students are dated, many of them being performed in the late 'fifties and early 'sixties. Noticeably, these populations have changed since this era. Further, there have been no reported studies dealing specifically with the effect of timed testing on the reading comprehension scores of developmental students, an identifiable segment of the total student population whose measured abilities and attitudes may be appreciably different from and are presumably not as efficient as those studied some years ago. Therefore, this study purposefully directs attention to the experience of this increasingly visible population that confronts the reading comprehension test situation. More specifically, it focuses on the effect of timing on the scores and reading/testing behaviors of developmental students.

THE STUDY

To investigate the effect of timed testing on developmental students' scores and to develop data that might better reveal students' reading and testing behaviors, three tests were administered to provide six variables.

First, the reading section of the Cooperative English Test, Form 1A, Coop 1A, was administered untimed so that students were allowed to complete all items on the test. This measure yielded a comprehension score; it also provided another variable: the amount of time each student used to complete the test. Given this testing environment, students had more time to engage in rauding (Carver 1984): to internally articulate each successive word in or part of a sentence in an attempt to deal with sometimes highly abstract material framed in

complicated sentences and sophisticated rhetoric (Carver 1976). Also, without time constraints, students were afforded time for scaffolding and bootstrapping (Feldman 1976; Smith 1985): they were allowed time to erect external and internal schematic frameworks to deal with concepts, abstractions, and allusions that might not easily fit into their insufficient schemata (Breiter 1985; Kaufman, Randlett and Price 1986; Lee 1984). Finally, removing time constraints helped reduce the anxiety that produces test-taking behaviors which can result in aberrant scores (Wark and Bennett 1981). Accordingly, this untimed testing situation was calculated to more nearly simulate a study session in which students endeavor to learn from texts that are substantially beyond their reading level but perhaps not entirely beyond their ability to raud (Rodriguez 1985).

Second, the reading section of the Cooperative English Test, Form 1B (Coop 1B) was administered under timed conditions and per testing manual directions. This measure yielded comprehension scores; additionally, students' answer sheets were surveyed to ascertain the number of comprehension items that each student had attempted. In this testing situation, students were subjected to speeded testing conditions that are typical and, perhaps, are regarded as normal. Students taking Coop 1B most probably did not have enough time to engage in rauding, bootstrapping, scaffolding or any other supportive strategies.

Finally, the Degrees of Reading Power, PA-2 test (DRP) was administered untimed and according to manual directions. This measure yielded comprehension scores: also, each student's completion time was recorded. While taking the DRP, students again had time to engage in the supportive reading behaviors that they could use on the Coop 1A. However, the DRP, which is designed as an untimed instrument, employs passages of increasingly more difficult reading material so that subjects are initiated to passages of greater difficulty. Thus, passages with a wider range of difficulty are presented to students.

The three tests were administered at the following intervals. Coop 1A Untimed was administered, and two days later the next class meeting Coop 1B Timed was given. Four class meetings hence, the DRP Untimed was administered. The data derived from these three testing instruments were organized and treated to produce the correlational studies and to present other comparisons that provide the empirical basis for this study.

The population studied consisted of ninety-five students in four day-time community college developmental reading classes conducted during the fall 1984 semester. These students quite clearly conformed to the developmental definition found in the literature (Cross 1976; Boylan 1985; Promises to Keep 1983). Their reading placement scores could be loosely interpreted as representing reading levels between the fifth and the ninth grades. Only two of these students had completed a college preparatory high school program. Fifty percent were receiving financial aid and/or using support services offered by a compensatory program (EOPS). Ninety percent held full- or part-time jobs, and eighty-five percent were enrolled in other college courses. If research findings based on contiguous community college populations can be generalized, sixty-five percent of these developmental students were regular (daily or weekly) users of alcohol, and thirty-six percent were regular users of other psychoactive drugs (marijuana, amphetamines, cocaine) in descending order of use or abuse (Skager and Maddahian 1984).

FINDINGS

Does timing affect students' performance on reading comprehension tests?

Three comparisons are presented to indicate the effects of timing on reading comprehension test scores of developmental students.

First, students' scores on Coop 1A Untimed were compared to their scores on the DRP Untimed to yield a correlation of r = .48 (Table 1). The fairly strong positive relationship between these two untimed testing situations obtains despite the fact that the tests' formats differ appreciably, and one test contains passages and items that present a greater range of difficulty.

To reveal the direct effect of timing on the results of an otherwise equivalent test, students' scores on the Coop 1A Untimed were compared with Coop 1B Timed to produce a correlation of r = .18 (Table 2). Since timing was the only variable distinguishing these testing conditions, timing or the speed set that it occasions appears to be responsible for the low correlation of the two test forms, which are otherwise correlated respectably at .66 and .72, depending on the publisher's norm table consulted (Coop Manual Technical Report 1960, p. 21).

A third comparison again suggests that timing or time pressure affects developmental students' scores on reading comprehension tests. Students' scores on the Coop 1B Timed were compared to their scores on the DRP Untimed to reveal virtually no relationship: r = .054 (Table 3).

The timed/untimed conditions of test administration, then, apparently account for the differences in students' scores and also the two low correlations stated immediately above.

How do students' speed of response behaviors affect scores on reading comprehension tests?

Speed of response can be measured in two ways. On a timed test, it can be measured by the number of responses a subject records during the timing period. On an untimed instrument, it can be derived from the amount of time it takes a subject to complete all items on the test. The term speed of response is not equated with the term speed of comprehension since the latter does not purposefully take into account

random responding. Therefore, in this paper, the more restrictive term speed of response will be employed.

To learn the effect of speed of response on a timed test, the number of student responses on Coop 1B Timed were compared with scores on that test to yield a strong correlation of r = .61 (Table 4). This high correspondence is normal expectation, for the greater the number of responses registered, the greater the chance for scoring unless the test is scored with correction for chance success. Further, the condition that a more rapid rate of response is responsible for generating higher scores on time-critical reading comprehension tests is consistent with similar or identical research performed on broader student populations (Davis 1962; Stetson 1982).

To learn the effect of speed of response on untimed reading comprehension test scores, two relational sub-studies were performed. The number of minutes students used to complete Coop 1A Untimed were compared to their scores on that test to produce a correlation of r = -.06 (table 5). Again, the number of minutes students used to complete the DRP Untimed were correlated with their scores on that test to yield r = .003 (Table 6). It is apparent, then, that speed of response as measured on untimed tests is not a factor that influences students' scores on the tests themselves.

Is speed of response a transcendent behavior that can be related to scores on other reading comprehension tests?

To learn whether speed of response persists to affect other testing situations, four sub-studies were conducted.

First, students' speed of response on Coop 1A Untimed was compared to their scores on Coop 1B Timed to reveal r = -.19 (Table 7). Then students' speed of response on the DRP Untimed was compared to scores on Coop 1B Timed to show a correlation of r = -.10 (Table 8).

Thus, there is negligible relationship between the rate at which students respond on tests that are untimed and students' scores on a time-critical test.

Next, students' speed of response on Coop 1B Timed was compared to their scores on the DRP Untimed: r = -.22 (Table 9). Then students' speed of response on Coop 1B Timed was compared to their scores on Coop 1A Untimed: r = -.17 (Table 10). Again, there is negligible relationship between students' rate of response on a timed test when compared to scores on untimed tests.

Therefore, while students' speed of response on a time-critical test is positively related to their scores on that test, there appears to be nothing inherent or residual in their speed of response behaviors that influences scores on other tests of reading comprehension whether the tests be time-critical or untimed.

How do developmental students' speed of response behaviors compare with those of a more typical student population?

To compare the response behaviors of developmental students under timed testing conditions with the experience of a wider student population, the response experience of this developmental group on Coop 1B Timed was compared to publisher's findings concerning its experimental population (Coop Manual Technical Report 1960, pp. 21 - 22).

FIGURE A

	Developmental Students	Coop Experimental Population
Students Responding to Fewer than Thirty Items	55%	10%
Students Responding to Thirty Items or More	43%	90%
Students Responding to the Last Test Item	2%	15%

The obvious differences illustrate the rate behavior discrepancy between developmental students and a broader, more typical population used by the publisher. Perhaps predictably, developmental students' speed of response behaviors are markedly slower than those of the general student population, especially at the extremes of the response continuum. This comparison is also consistent with developmental students' comparatively low performance on Coop 1B Timed.

What is the effect of students' speed of response on their accuracy of response?

Thus far this paper has investigated the effects of timed and untimed testing on students' test scores and has shown how students' rate of response influences scores on both types of instruments. It has also noted that a reading test score on a time-critical test may be sensitive to rate-of-response and chance-success behaviors. Now attention will be directed to another matter: the effect of timing on students' accuracy of response to test items.

Unlike the raw score measure on a time-critical test, the accuracy of response measure is not susceptible to speed of response behaviors. This condition can be observed by comparing the number of items attempted on a timed test (Coop 1B Timed) with the percentage of accuracy score on that test: $r = -.09$ (Table 11). Thus, students' accuracy of response is unaffected by speed of response; contrarily, and as previously noted, students' speed of response is correlated ($r = .61$) with their raw scores on Coop 1B Timed (Table 4). Since the percentage of test items scored correctly is a measure that is indifferent to students' speed of response behaviors, accuracy of response is a factor that directs attention.

Accordingly, the following comparison is presented to show the effect of timing on students' accuracy of response.

FIGURE B

Test	# of Items on Test	Avg. # of Items Attempted	Mean Score	Avg. Accuracy of Response
Coop 1B Timed	60	31	10.77	35%
Coop 1A Untimed	60	58	30.26	55%
DRP Untimed	77	77	65.58	85%

Taking the test under the prescribed timing conditions, students scored at thirty-five percent accuracy on Coop 1B Timed. Without time constraints and given the chance to answer all test items on an otherwise equivalent instrument (Coop 1A Untimed), students answered at fifty-five percent accuracy. And on an untimed test incorporating passages with a wider range of difficulty (DRP Untimed), students' accuracy of response reached eighty-five percent. The effect of timing, then, is that it diminishes students' accuracy of response (Carver 1984).

How does the accuracy of response of developmental students compare with the experience of a wider population?

Another comparison is presented to show how developmental students' rate of response is related to their accuracy of response.

Previous research has shown that, on a time-critical reading comprehension test, as students' response rates rise and accelerate, their accuracy of response diminishes (Botel 1951; Davis 1962; Flanagan 1939; Stetson 1982). The following comparison is presented to determine whether the experience of developmental students is comparable.

FIGURE C

	N	%	Accuracy of Response
Students Responding to Fewer than 30 Items	52	(55%)	35%
Students Responding to 30 or More Items	43	(45%)	35%

Developmental students responding to fewer than thirty items were responding below publisher's expectation (see Figure A). This group experienced thirty-five percent accuracy of response. Those responding to thirty or more items were performing at or above publisher's minimum expectation and scored at the same degree of accuracy. Thus developmental students' measures of accuracy of response do not correspond to findings of other researchers dealing with non-developmental populations.

What is the overall accuracy of response performance of developmental students on a timed test?

Two additional findings with respect to timed testing and accuracy of response may be most pertinent. First, of the ninety-five students who took Coop 1B Timed, twenty-one students scored at twenty-five percent accuracy or lower. Which is to say that twenty-two percent of the students scored at or below chance on a timed reading comprehension test. Finally, the accuracy of response for the entire developmental population (N=95) was thirty-five percent on Coop 1B Timed (see Figure B). Therefore, on a time-critical reading comprehension test the average student scored at ten percent above chance.

SUMMARY AND CONCLUSIONS

Early on, this study revealed that the relationship between developmental students' scores on untimed reading tests is positive (Table 1). However, the correlations between scores on a time-critical reading comprehension test and two untimed reading tests are negligible, these low correspondences being attributable to the timing variable (Tables 2 and 3). Timing or the speed set it elicits, then, not only affects scores but also the manner in which a test measures.

It was learned that developmental students' scores on a time-critical test are profoundly affected by students' speed of response

(Tables 4, 5, and 6). Although timed tests provide opportunity for chance-success scoring, they purport to measure comprehension. But if scores on timed tests are indeed attributable to information or knowledge derived from the reading passages presented and are not affected by the artifact of completing more items then

- scores on timed tests would be positively correlated with scores on untimed tests, which reflect a higher degree of accuracy of response;
- the number of items attempted on an untimed test would correlate positively with accuracy of response on that test;
- students' average accuracy of response scores on a timed test would exceed thirty-five percent or ten percent above random.

However, as we have seen, none of these conditions obtains (see Tables 2 and 3; 11; and Figure C, respectively). In fact, not only do time-critical tests reward rate behaviors that are not clearly associated with comprehension, but they also penalize students whose slowness may occasion accuracy.

Next, if there is some germinative factor latent or inherent in speed of response that persists to influence the comprehension testing conditions of another instrument or instrumental application, then

- students' rate of response under the less pressing circumstances of an untimed test would correlate with their scores on a timed test

and/or

- students' speed of response under the pressured conditions of timed testing would correlate with students' scores on untimed tests.

Again, these comparisons indicate that neither of the above conditions prevails (see Tables 7 and 8, and 9 and 10 respectively). There is no evidence, then, that any identifiable rate patterns endure to exert

influence beyond the confines of their own instrumentation. Indeed, it would appear that there is no quality or value in speed of response that can be related to reading comprehension.

It is apparent that time-critical reading comprehension tests, which critics view as unsuitable for student populations as a whole, are even less appropriate for developmental students whose rate behaviors are comparatively retarded. In fact, derliberative readers or plodding rauders would necessarily have to be exceedingly accurate responders to score creditably on a time-critical reading comprehension test. Again, developmental students, whose rate behaviors and whose measured accuracy of response remain relatively stable or constant on time-critical tests (see Figure C) either cannot or do not take advantage of the artifactal influence of speed of response on scores. Thus, they are particularly vulnerable to the measurement bias of timed testing and may become a score casualty because of the accident of habitual deliberative reading.

Time-critical reading comprehension tests suffer from a design flaw because their measures are obtained from samples that occasion speed set. Subsequently, these tests do not simulate real-world conditions. Accordingly, timed testing does not address itself to directly or precisely assessing the basic abilities which facilitate a students' functioning in a situation that involves learning from text (McDonald 1966; Tuinman 1971). While ignoring testing conditions that might provide a representative sampling of reading comprehension behaviors, time-critical tests reward rate behaviors that, as far as can be determined, are inconsequential.

Finally, speeded comprehension tests are inappropriate for reasons that go beyond considerations of test validity. First, they encourage behaviors of testmanship — guessing, wild guessing, and random responding — whose sole function is score escalation. Next, they discourage the deliberative reading that even the most

advanced readers typically employ to understand challenging material (Carver 1985; Feldman 1976). Finally, they are responsible for contributing to the perhaps deserved loss in credibility that our testing and placement programs have suffered; for a test that has a hidden agenda is guilty of duplicity, and the educational industry can ill-afford perpetuating doubt about its integrity.

TABLES

TABLE 1

Score, DRP Untimed	M = 65.58	SD = 6.55	r = .48
Score, Coop 1A Untimed	M = 29.55	SD = 5.68	(N = 81)

TABLE 2

Score, Coop 1A Untimed	M = 30.29	SD = 5.02	r = .18
Score, Coop 1B Timed	M = 10.81	SD = 4.93	(N = 95)

TABLE 3

Score, DRP Untimed	M = 65.58	SD = 6.54	r = .054
Score, Coop 1B Timed	M = 10.91	SD = 6.89	(N = 81)

TABLE 4

Score, Coop 1B Timed	M = 10.87	SD = 4.93	r = .61
Number of Items Attempted, Coop 1B Timed	M = 30.01	SD = 9.85	(N = 95)

TABLE 5

Score, Coop 1A Untimed	M = 30.29	SD = 4.99	r = -.06
Number of Minutes Used to Complete Coop 1A Untimed	M = 87.97	SD = 14.67	(N = 95)

TABLE 6

Score, DRP Untimed	M = 65.58	SD = 6.55	r = .003
Number of Minutes Used to Complete DRP Untimed	M = 84.80	SD = 16.19	(N = 81)

TABLE 7

Score, Coop 1B Timed	M = 10.87	SD = 4.93	r = -.19
Number of Minutes Used to Complete Coop 1A Untimed	M = 87.97	SD = 14.67	(N = 95)

TABLE 8

Number of Minutes Used to Complete DRP Untimed	M = 84.80	SD = 16.70	r = -.10 (N = 81)
Score, Coop 1B Timed	M = 10.91	SD = 4.99	

TABLE 9

Number of Items Attempted, Coop 1B Timed	M = 30.98	SD = 10.02	r = -.22
Score, DRP Untimed	M = 65.76	SD = 6.62	(N = 81)

TABLE 10

Score, Coop 1A Untimed	M = 30.29	SD = 4.84	r = -.17
Number of Items Attempted, Coop 1B Timed	M = 30.01	SD = 9.85	(N = 95)

TABLE 11

Number of Items Attempted, Coop 1B Timed	M = 30.01	SD = 9.45	r = -.09 (N = 95)
Percentage of Accuracy, Coop 1B Timed	M = 35.13	SD = 11.97	

References

Barton, Norma. Standardized reading testing in college: some cautions, in R. Sugimoto. Ed., Proceedings of the 8th Annual Conference of the Western College Reading Association, 1975, 31-4.

Boylan, Hunter. Underpreparedness: a reinterpretation of the problem and some recommendations for resolution. Position Paper of the National Association for Developmental Education (1985) P.O. Box 60227, Chicago, IL 60660.

Breiter, Carl. Toward a solution of the learning paradox. Review of Educational Research, 55, 2, 1985, 201-226.

Carey, Robert. Empirical vs. natural research? Reading Research Quarterly, 16, 3, 1980, 412-15.

Carver, Ron. Measuring readability and reading ability in order to predict comprehension, in W. Miller and G. McNinch, Eds., 25th Yearbook of the National Reading Conference, 1976, 293-97.

Carver, Ron. Toward a theory of reading comprehension. Reading Research Quarterly, 13, 1, 1977-78, 8-61.

Carver, Ron. Rauding theory predictions of amount comprehended under different purposes and speed reading conditions. Reading Research Quarterly, 29, 2, 1984, 205-18.

Carver, Ron. How good are some of the world's best readers? Reading Research Quarterly, 20, 4, 1985, 389-419.

Cooperative english tests, manual for interpreting scores, Educational Testing services: New York, 1960.

Cross, K. Patricia. Accent on learning: improving instruction and reshaping the curriculum. San Francisco: Jossey-Bass, 1976.

Davis, Frederick. Measurement of improvement in reading skill courses, in E. Bliesmer and R. Staiger, Eds., 11th Yearbook of the National Reading Conference, 1962, 30-40.

Degrees of reading power. College Entrance Examination Board: New York, 1981.

Flanagan, J. C. A study of the effect of comprehension of varying speeds of reading, Research on the Foundations of American Education, Official Report of the American Educational Research Association, 1939, 47-50.

Kahn, Gloria. The effects of testwiseness on reading scores and on test anxiety, in W. Miller and G. McNinch, Eds., 25th Yearbook of the National Reading Conference, 1976, 113.

Kaufman, Nancy; Randlett, Alice, and Price, Jay. Awareness of the use of comprehension strategies in good and poor college readers. Reading Psychology, 6, 1986, 1-11.

Kingston, Albert J. Cautions regarding the standardized reading test. 4th Yearbook of the Southwest Reading Conference, 1955, 11-16.

Langer, Gary; Wark, David; and Johnson, Sue. Test-wiseness in objective tests, in P. L. Nacke, Ed., 22nd Yearbook of the National Reading Conference, 1973, I, 280-87.

Lee, Linda Ann. A piagetian theory for improving college reading abilities, in D. Fleming, Ed., Journal of College Reading and Learning, 1984, 17, 150-156.

McDonald, Arthur. Measuring reading performance, in G. Schick and M. M. May, Eds., 15th Yearbook of the National Reading Conference, 1966, 216-221.

McDonald, Arthur. Reading versatility twelve years later, in F. P. Greene, Ed., 20th Yearbook of the National Reading Conference, 1971, 168-173.

Preston, R. and Botel M. Reading comprehension tested under timed and untimed conditions. School and Society, 74, August 1951, 71.

Promises to keep: remedial education in california's public colleges and universities. California Postsecondary Education Commission Report 83-2: Sacramento, 1983.

Rankin, Earl. The relationship between reading rate and comprehension, in E. Bliesmer, and R. Staiger, Eds., 11th Yearbook of the National Reading Conference, 1962, 1-5.

Rodriguez, Joan Hughes. When reading less can mean understanding more. Journal of Reading, 28, 8, May 1985, 701-705.

Rounds, Jeanine. Assessment and placement in language arts courses: three major tests, in D. Fleming, Ed., Journal of College Reading and Learning, 1984, 17, 94-100.

Skager, Rodney and Maddahian, Ebrahim. A survey of the substance use and related factors among community college students in the county of orange, fall 1983. Graduate School of Education, UCLA, 1984.

Stetson, Elton. Reading tests don't cheat, do they? Journal of Reading, 25, 7, April 1982, 634-9.

Tillman, Chester. Readability and other factors in college reading tests: a critique of the diagnostic reading test, the nelson-denny reading test, and the mcgraw-hill basic reading test, in D. Pearson and J. Hansen, eds., 26th Yearbook of the National Reading Conference, 1977, 253-258.

Wark, David. Twenty-five years of research on adult reading, in E. Thurston and L. Hafner, Eds., 14th Yearbook of the National Reading Conference, 1965, 218-223.

Wark, David and Bennett, Michael. The measurement of test anxiety in a reading class. Reading World, 20, 3, March 1981, 215-222.