THE EFFECT OF TEST TIME LIMITS ON PERFORMANCE FOR CONTROLLED AMOUNTS OF STUDY

BRIAN DEAN WILBUR

Copy 2

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

NL-339 (3/77)

THE EFFECT OF TEST TIME LIMITS ON PERFORMANCE FOR CONTROLLED AMOUNTS OF STUDY

BY

Brian Dean Wilbur, A.B., B.Ed.

A Thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Education

Department of Educational Psychology
Memorial University of Newfoundland

November 1975

St. John's Newfoundland

ABSTRACT

The study investigated the relationship between imposed test time limits and students' test performance for controlled amounts of student exposure to examined information.

Nine classes of grade eight students within the Avalon North Integrated School District were randomly selected and randomly assigned to one of the nine possible treatment conditions.

The material relevant to the achievement measure was contained in a ten minute audio-filmstrip presentation and was supplemented with a summary to be read by the student. The three levels of the exposure to information variable included one and two exposures to the examined information, and a no exposure condition that served as a control condition. The three levels of the time limit variable, short, medium and untimed, were established by a pretest of the items selected for the final achievement measure.

The results indicated a significant positive correlation between correct responses and omitted items. Apparently, the more information Ss had about the tested material, the more they omitted making responses about which they were not certain. Conversely, the less information Ss had about the tested material, the less they omitted making responses about which they were uncertain. The results also indicated a significant negative correlation between available time and omitted responses. As time increased, all Ss seemed to construct responses rather than omit items. The exposure to material variable produced a significant main

effect. Both exposure treatments were significantly different from the control condition but they were not significantly different from each other. For the number of correct responses there were no significant differences for the various time conditions.

ACKNOWLEDGMENTS

The writer wishes to thank Bryan Hartman, Supervisor, for his advice, time and understanding throughout this thesis. Appreciation is also extended to Dr. Terrance Boak and Dr. Wayne Nesbit for their constructive criticism. Acknowledgment is given to the receipt of a fellowship from Memorial University of Newfoundland.

Special gratitude is due my wife, Jean, for her encouragement, patience and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

In many schools students above the elementary level may be subjected to a considerable amount of formal testing, often as much as one formal test every three school days, not including subject quizzes or standardized tests. This amount of testing takes time. The student must take time away from his class study to take the tests. The teacher must interrupt the instructional plan to find time to prepare, administer and score these instruments. Since the tests take time, they should afford the student an educational benefit in return for the time spent taking examinations. Consequently, any factors that reduce test validity should be identified and corrected. One common testing practice that may be a source of test invalidity is the practice of imposing arbitrarily determined time limits on examinations. This study will investigate the relationship between imposed test time limits and students' test performance when the amount of exposure the students have to the examination material is controlled.

Unfortunately, the setting of time limits during testing sessions is often effected only for administrative convenience. Cronbach (1949) noted:

> Timed tests are often needed for administrative convenience, but when speed becomes a major element in determining a person's score, the score is likely not to represent his attainment accurately. Speed is a legitimate element in achievement tests only when speed is an objective of the course (p. 268).

Glaser (1963) stated the scores obtained from achievement tests provide two kinds of information: the degree to which the student has attained criterion performance and the relative ordering of individuals in terms of their test performance. In theory, and generally in practice, most school testing has the former as an objective. The latter is carried out at most academic levels often only to satisfy administrative demands. The relative ordering of individuals may be accomplished by means of either a power test or a speed test. A power test includes questions that range from simple to complex. On these measures it is expected that a student will correctly respond only to those questions that do not exceed his ability. A speed test is based upon the assumption that response speed is a direct function of ability. These measures impose time limits that will allow students to complete different numbers of test items in direct relation to their ability. Often a combination method is arrived at by combining both power and speed methods to determine relative performance distributions.

The question of the roles that speed and power play in testing has long been debated. Around the first quarter of this century many articles were written concerning the effect of speeded responding upon test performance, and whether speed or power was the most valid method of determining relative degrees of task-related achievement. Early investigators found positive correlations between speed and power tests. Therefore it was concluded that both methods of testing were valid indicators of achievement (Freeman, 1928; Longstaff and Porter, 1928; Ruch, 1923).

Questions about the relationship between response speed and achievement level have also been debated. Wechsler (1949, 1955, 1974),

by the inclusion of the coding subtest of his intelligence scale, directly implied that the response latency associated with very simple learning was significantly related to intelligence. The inclusion of bonus marks for quick responses on the picture arrangement, block design and object assembly subtests give further support to the observation that Wechsler's measure includes speed as a valid part of task achievement. Similarly, many of the standardized testing measures used in the schools contain an element of speed. For this reason, the tests might be generally considered to be a mixture of power and speed with the emphasis on power. Since speed is a part of many standardized achievement measures used, Stafford (1971) suggested that a new indicator be added to achievement measures. He proposed a speededness quotient derived by dividing the total number of unattempted items by the sum total of the number of questions answered incorrectly, the number attempted but omitted and the number unattempted. A pure speed test would have a speededness quotient of 100% while a pure power test would have a quotient of 0%.

Several studies have indicated that speed is a valid index of task achievement. Knapp (1960) administered a cross-cultural intelligence test under speed and power conditions to a Mexican and American sample of adult males. The results indicated that Ss tended to maintain their relative positions in the group under both administrations. In a different vein, Briggs and Johnson (1942) recorded the time students took to complete a final exam. They divided the final exam papers into three groups according to the time when the students finished the exam. The students who submitted their papers in the first and last thirds of the distribution had higher achievement scores than those who submitted

their papers in the middle third.

In contrast, other studies have concluded that speed is not a valid index of achievement (Boag and Neild, 1962; Daly and Stahmann, 1968; Kendall, 1964; Michael and Michael, 1969; Mollenkopt, 1950, 1960; Terranova, 1972; Wesman, 1960; Yates, 1966a, b). These studies indicate that speed acts as a detriment to achievement assessment.

Mollenkopf (1950), investigating item placement under speed and power conditions, found that:

> Those who more often change their responses in any way when given added time tend to be individuals who work fast, and who get higher scores under the speed conditions. The changes tend to better the scores for the faster individuals more than for the slow ones, and tend to be made in the right direction more often by high-scoring than by low-scoring students (p. 297).

Boag and Neild (1962), using a vocabulary test, found that the average high school student increased his score and often changed his relative standing when additional test time was given.

Yates (1966a, b), in a series of studies, administered the Progressive Matrices and an arithmetic test. Both tests were administered with an announced initial time limit which was extended at the end of the test. Yates argued that when the Progressive Matrices was administered with conventional time limits, it confounded the level and speed of intellectual performance and seriously underestimated the intellectual capacity of a small group of very capable students. This finding was supported by the arithmetic test which was constructed in cycles. The cycles were repeating sets of questions which were scored as independent units. Therefore the number of sets completed did not affect the results so that a slow response speed was not a handicap.

Mollenkopf (1960) noted that the same material given under the

conditions of speed and power may not be measuring the same behavior. The assessment may become an expression of speed rather than a reflection of knowledge.

The decrement in test performance produced by the setting of time limits may be of considerable magnitude. Daly and Stahmann (1968) concluded that the effect on scores of additional testing time for university students on the Cooperative English Test was of the same magnitude as the remedial English course to which they were assigned.

Investigations of the effect of an unlimited time condition have found that the amount of time a student takes to complete a test was not related to his score. Michael and Michael (1969) reported that with one exception out of ten groups, there was no significant relationship between the amount of time a university student took for various psychology tests and his level of performance. Similarly, Terranova (1972) administered an ability test to randomly selected fifth, seventh, ninth, and eleventh grade students. He concluded that there was no significant relationship between test scores and the amount of time taken to complete the test. Kendall (1964) investigated aptitude testing noting six different time measures on the answer sheets. He concluded that there were no significant relationships between ability level and the location of maximal time limits. In a review of the literature Wesman (1960) concluded that it was primarily knowing more, not thinking faster, that distinguished those who scored high from those who scored low. It was only on measures of skill requiring very little understanding that speed showed a direct relationship with performance.

Another factor found to be related to test time and achievement level was test taking skill. Bennett and Doppelt (1956) compared item

difficulty with response speed. They found that the less able subjects worked at about the same rate with easy and difficult materials while the more able subjects responded rapidly to easy items and slowly to difficult items. This result implied that the more able subjects demonstrated a greater degree of efficiency thereby increasing the number of items which could be attempted. Dubin, Osburn, and Winick (1969) hypothesized that extra pretest practice, extra testing time, and combinations of both would improve the mental ability test performance of Negroes more than Caucasians. While support for their differential hypothesis was obtained, both racial groups significantly improved their performance to a similar degree as the standards became more lenient and practice increased.

When a time limit for a teacher-made test is set, factors, such as: previous educational experience of the students, amount of exposure to the material on which the assessment is based, and amounts of time required by the class on previous tests are often considered. Unfortunately, little evidence is available to determine the relative importance of these and possibly other factors, in relation to an assessment of the influence of imposed time limits on student performance. More specifically, evidence relating to the relationships among the amount of study, the amount of time allowed during testing and test performance was not available in the measurement literature.

In addition to teacher-made tests, the nature of most standardized achievement tests preclude time adjustments for individual differences. Therefore, on a standardized achievement test the effect of time limitations and previous exposure to material may again combine and produce a confounding effect. This possible confound would be of particular

interest to the counselor, as one who generally is responsible for the selection, administration and interpretation of a variety of standardized achievement measures. Since counselors generally use the results of both teacher-made and standardized tests to assist students to make both educational and vocational decisions, the effect of time limits on test performance would be information of considerable assistance to both the student and the counselor.

Specific to the proposed investigation the following hypotheses were postulated. First; in that repetition has been found to be positively correlated with retention (Thorndike, 1932), it was anticipated that the Ss level of retention would be controlled by the number of exposures to the testable material. Consequently, the study variable was operationally defined as the number of exposures to the testable material and it was hypothesized that there is a direct positive relationship between the amount of study and the number of correct responses.

Second, as noted in the above review of literature, the relationship between test time limits and performance is a contentious issue. Therefore, a null-form research hypothesis was postulated that there was no significant relationship between total response latency and the number of correct responses.

Third, the possibility of an interaction between the number of exposures and the length of the response latency was unknown; therefore, a null-form research hypothesis was again postulated. It was hypothesized that there was no significant relationship between the number of correct responses and response latency for any exposure condition.

CHAPTER II

METHOD

Design

The design was a randomized control-group post-test only (see Figure 1). This design was selected to eliminate the possible interaction effect of pretesting. If pretesting was used, subjects would be immediately alerted to the criteria selected for the achievement testing.

Subjects

Nine classes of grade eight students within the Avalon North Integrated School District were randomly selected from a population of seventeen classes. Grade eight was the most appropriate grade for this material since the program of studies included European geography. Each of the nine classes was randomly assigned to one of the nine possible treatments within the 3 X 3 factorial design. To establish equal numbers of subjects within cells, 22 subjects were randomly selected from each cell of the design for data analysis.

Procedure

The material relevant to the achievement measure was contained in a ten minute audio-filmstrip presentation, Europe's Common Market: Problems and Prospects. The material was supplemented with a summary (Appendix A) to be read by the student. The presentation was viewed by a class of students either once or twice depending upon the treatment procedure to which they were randomly assigned. A control group did not

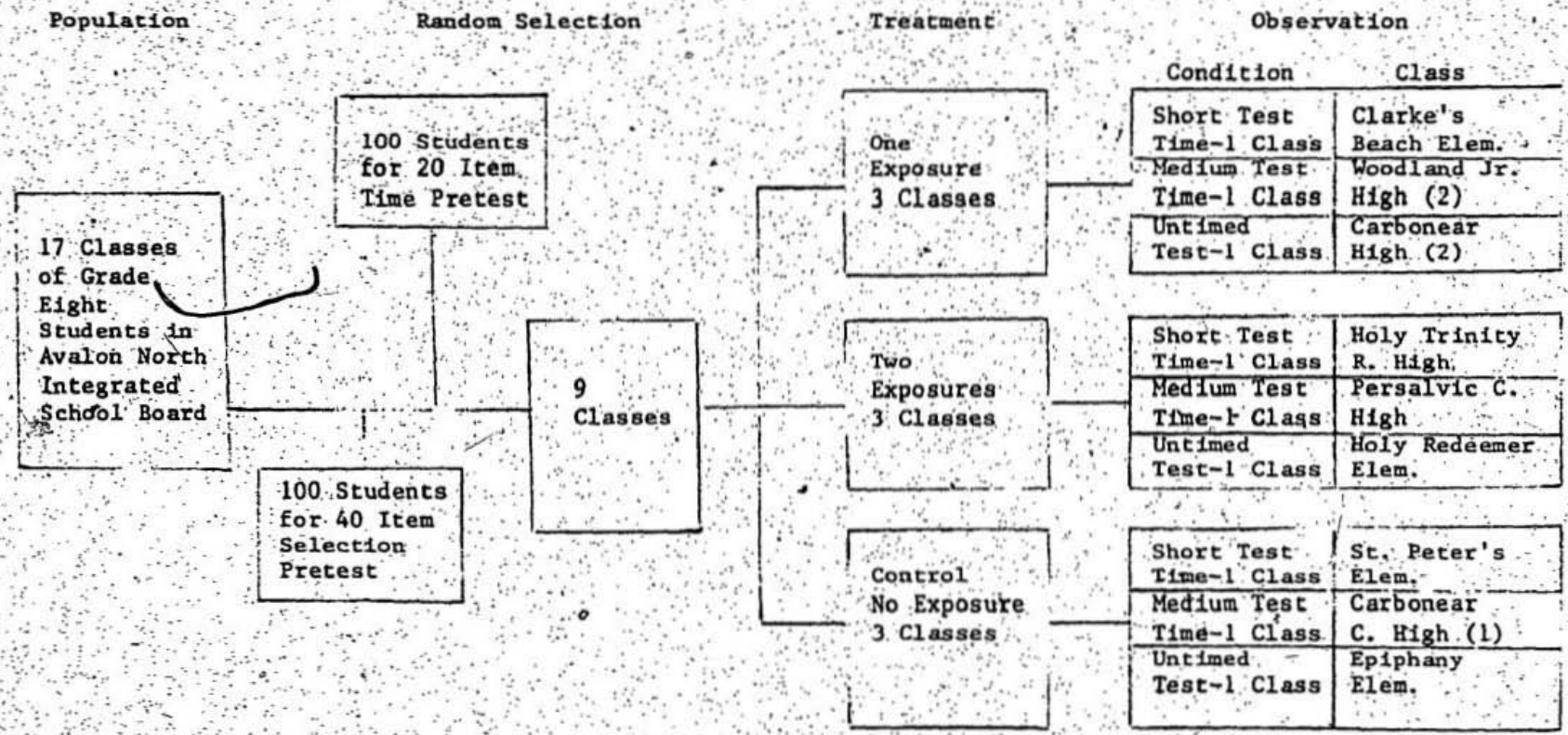

Figure 1. Experimental procedures chart

see the presentation.

After exposure to the presentation, treatment subjects were asked to respond to a set of completion questions concerning the film. The amount of time the subjects had to respond to the questions was announced to them prior to testing. The time intervals assigned were the short, medium and untimed intervals that were determined by the results of a pretest described below.

Pilot Studies

The achievement measure of completion questions required the students to respond with a short written answer after one presentation of the material (Appendix B).

The results of the forty item pretest determined the composition of the final items used for the assessment. The pretest for item selection had a test mean of 6.61 and a variance of 29.99. The Kuder-Richardson formula 20 reliability was 0.8548. The selection of items for the final test was based on two criteria: item difficulty and biserial correlation. The item difficulty indicated the percentage of the pretest sample who responded correctly to the item. Since the questions in the pretest proved to be very difficult, with items ranging in difficulty from .0 to .615, it was necessary to select the easiest items so that a maximum number of Ss would have an opportunity to respond correctly. The biserial correlation is the correlation between the discrete distribution of right and wrong responses for each item and the continuous distribution of total correct responses for the entire test. The items with the highest positive correlations with the total test were selected for the final test (see Table 1). The Kuder-Richardson formula 20 reliability was 0.8690 for the final test.

Table 1

Pretest Difficulty Levels and Biserial Correlations of Items Selected for the Test

Test Item Number	Pretest Item Number	Difficulty	Biserial Correlation
1	1	0.125	0.690
2	3	0.167	0.420
3	4	0.500	0.374
4	7	0.188	0.961
5	11	0.177	0.707
6	12	0.260	0.502
7	14	0.167	0.991
8	15	0.458	0.374
9	16	0.354	0.594
10	17	0.302	0.819
11	18	0.365	0.716
12	19	0.229	0.637
13	21	0.250	0.666
14	23	0.271	0.552
15	24	0.615	0.397
16	25	0.333	0.582
17	31	0.260	0.783
18	32	0.260	0.437
19	36	0.281	0.617
20	38	0.135	0.805

The difficulty levels of a sample of forty questions about the presentation were determined by a pretest of ninety students. A subtest of twenty questions of this sample was selected such that the mean level of difficulty of the final achievement measure was .285.

The time conditions were determined by a pretest during which the amount of time a sample of one hundred and eight students took to respond after one presentation was individually recorded. A condition of short (1/3 of the total time required for all students to complete the test, 291 seconds) and medium (2/3 of the total time to completion, 582 seconds) was established. Of the pretest sample no students completed the test within the short time condition, forty-eight percent within the medium condition. In the untimed condition the students were given as much time as they required to complete the task and were specifically told this before beginning. The longest time required in this condition was 1,657 seconds.

CHAPTER III

RESULTS

The results of the final twenty item test indicated the variables right, wrong and omit had means of 5.03, 13.47 and 1.50 respectively. The correlations between these variables were computed (right versus wrong $r = -0.872$; $p < .01 = .254$: right versus omit $r = -0.134$; $p < .05 = .195$: wrong versus omit $r = -0.368$; $p < .01 = .254$) (see Table 2).

Table 2

Test Correlations for Right, Wrong and Omitted Items

	Right	Wrong	Omit
Right		-0.872	-0.134
Wrong			-0.368
Omit			

The data were analyzed by means of a 3 (exposures) X 3 (time limits) fixed effects model analysis of variance for equal N's. The probability level for all tests of significance was set at $\alpha = .05$. Multiple comparisons between means were tested by the Newman-Keuls procedure of magnitude ordered contrasts.

The ANOVA for the right responses indicated significant differences for the exposure variable (see Table 3). The Newman-Keuls, indi-

cated the no exposure condition was significantly different from both one and two exposure conditions at the .01 level. One exposure was not significantly different from two exposures. The analysis for right responses indicated no significant difference for the time variable (see Table 4). The interaction between the variables was not significant for the right responses (see Figure 2).

The ANOVA for the wrong responses indicated significant differences within the exposure variable (see Table 5). The Newman-Keuls comparisons indicated the conditions of no exposures was significantly different at the .01 level from that of both the one and two exposure conditions. The analysis for wrong responses also indicated a significant difference in the main effect for the time variable (see Table 6). The short time limit and the untimed condition were significantly different at the .01 level. The differences between the medium time limit and both the short and untimed conditions were not significant. In addition, the interaction between the variables, exposure and time, for the wrong responses was not significant (see Figure 3).

The ANOVA for the omitted responses indicated significant differences within the exposure variable (see Table 7). The Newman-Keuls indicated the omitted questions for the no exposure treatment were significantly different at the .05 level from one and two exposure treatments. One exposure was not significantly different from two exposures. This analysis also indicated a significant main effect for the time variable. The number of questions omitted was significantly different between the untimed and short time treatments at the .01 level and between the untimed and medium time treatments at the .05 level. The short time

Table 3

Summary Analysis of Variance for Right Responses

Source	Sum of Squares	df	F	p
A (exposures)	2059.756	2	64.068	0.000
B (time)	97.719	2	3.040	0.050
AB	123.338	4	1.918	0.109
Within	3038.135	189		

Table 4

Number of Right Responses for Each Variable

	Short Time	Medium Time	Untimed	Sum of Right Responses
1 Exposure	92.0	170.0	179.0	441.0
2 Exposures	162.0	183.0	170.0	515.0
No Exposure	10.0	16.0	5.0	31.0
Sum of Right Responses	264.0	369.0	354.0	

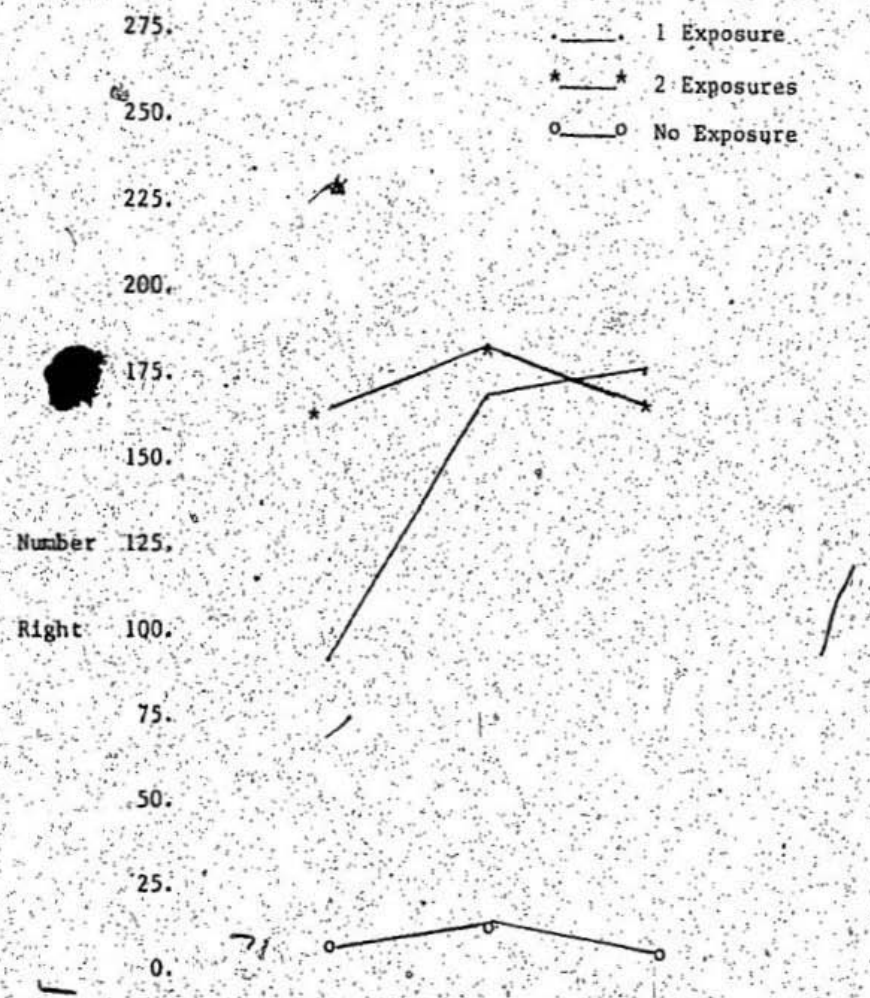

Figure 2. Number of right responses for each variable

Table 5

Summary Analysis of Variance for Wrong Responses

Source	Sum of Squares	df	F	p
A (exposures)	2660.723	2	80.563	0.000
B (time)	168.904	2	5.114	0.007
AB	86.630	4	1.312	0.267
Within	3121.000	189		

Table 6

Number of Wrong Responses for Each Variable

	Short Time	Medium Time	Untimed	Sum of Wrong Responses
1 Exposure	255.0	241.0	261.0	757.0
2 Exposures	189.0	222.0	270.0	681.0
No Exposures	373.0	420.0	435.0	1228.0
Sum of Wrong Responses	817.0	833.0	966.0	

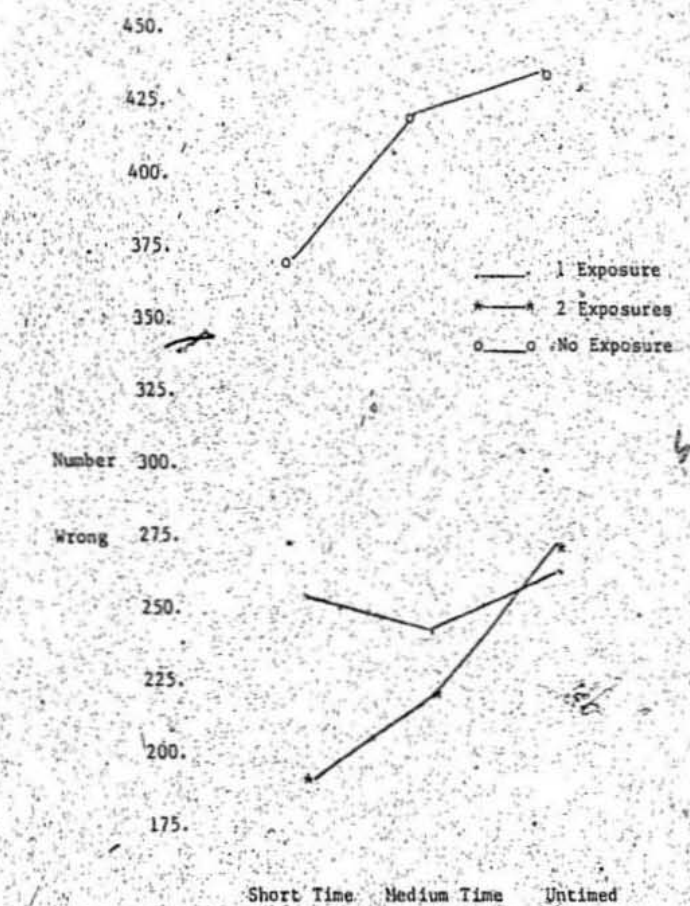

Figure 3. Number of wrong responses for each variable

condition was also significantly different from the medium condition at the .01 level of significance (see Table 8). The interaction between the exposure and time variables was not significant for the number of questions omitted (see Figure 4).

Table 7.

Summary Analysis of Variance
for Omitted Responses

Source	Sum of Squares	df	F	p
A (exposures)	33.302	2	3.317	0.038
B (time)	470.939	2	46.907	0.000
AB	18.484	4	0.921	0.453
Within	948.772	189		

Table 8

Number of Omitted Responses
for Each Variable

	Short Time	Medium Time	Untimed	Sum of Omitted Responses
1 Exposure	93.0	29.0	0.0	122.0
2 Exposures	89.0	25.0	0.0	114.0
No Exposure	57.0	4.0	0.0	61.0
Sum of Omitted Responses	239.0	58.0	0.0	

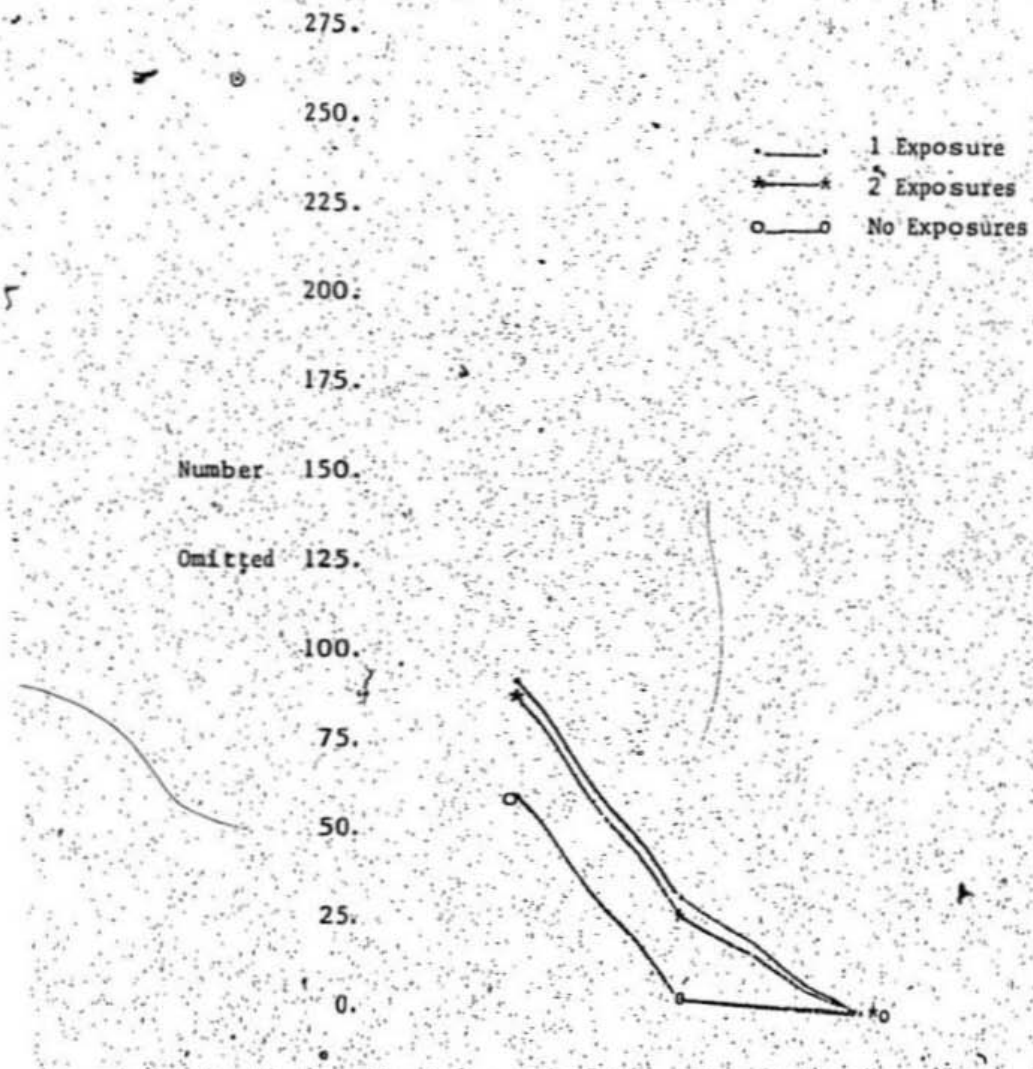

Figure 4. Number of omitted responses for each variable

CHAPTER IV

DISCUSSION

The study produced several results which are of interest to this investigator. First, the Pearson product-moment correlation between the wrong and omitted responses (-0.368) was unexpectedly significant. The correlation between wrongs and omits for short time conditions was -0.557, for medium time conditions -0.299 and for untimed conditions 0.0. This result suggests Ss in the short time condition tended to omit questions while Ss in the untimed condition tended to guess erroneously.

For the final test, the number of wrong responses for the medium time condition was not significantly different from the short and untimed conditions, but the wrong responses under the short time condition (817) were significantly lower than were those for the untimed condition (966). Since the number right does not change significantly between the time conditions, the number omitted changed in relation to the number wrong. The number omitted under the short time condition (239) was significantly different from the medium (58) and the untimed conditions (0.0). This prompts the conclusion that as Ss had more time, erroneous answers were often constructed so that the Ss could respond rather than omit the items. Therefore, speed may be a valid index of achievement in that unlimited response time, according to the results, only facilitates the construction of erroneous answers. If Ss were penalized for incorrect and/or omitted responses then the amount of time Ss have will affect the results. Unlimited time allows Ss to produce many erroneous responses and if Ss

were penalized for these then the unlimited condition will negatively affect the score. If omitted responses are a basis of penalization then short time limits will negatively affect the Ss results.

Second, the hypothesis, that there was a direct positive relationship between the amount of study and the number of correct responses, was partially supported by the data. The correct responses of the one and two exposure treatment conditions were significantly different from the no exposure treatment condition. The Ss exposed either once or twice to the material upon which the test was based scored significantly better than did the Ss who did not see the material but were tested. However, the repeated exposure failed to make a significant difference between the one and two exposure groups. The second exposure followed immediately after the first exposure and since the exposures were identical it would appear the close proximity of the exposures, allowed very little thoughtful consideration of the material, and/or produced a lack of attention to the second exposure; such that, the additional exposure was not an effective treatment.

For the number of correct responses there were no significant differences for the various time conditions. This result supports the second hypothesis that there is no significant relationship between the total response latency and the number of correct responses. The lack of relationship was in opposition to those who found speed a valid index of achievement (Briggs and Johnson, 1942; Wechsler, 1949, 1955, 1974) as well as those who concluded that speed is a negative index (Boag and Neild, 1962; Daly and Stahmann, 1968; Mollenkopf, 1960). The time spent responding to the questions did not differ significantly for the right responses. This implied the time allotted for the three conditions,

short, medium, and untimed, lacked the severity to create distinctive effects, or that the test was so difficult that Ss were able to quickly skip through the test, complete those questions with a high probability of responding correctly, then attempt to complete the remaining questions. With this strategy Ss would not benefit from additional time because the difficulty level and recall task would suggest that there would be a limited chance that subjects would be able to recall additional correct answers in the additional interval of time even though they do make additional responses.

The third hypothesis was supported by the data. There was no interaction between time and exposures to the material. However, this result may be partially due to the limited effectiveness of the exposure variable.

Future research on this topic should attempt to improve the effectiveness of the exposure variable by increasing exposures to material, varying types of exposures, and allowing various amounts of time between exposures. Such improvements are a prerequisite to a better understanding of the relationship between test time limits and study.

REFERENCES

REFERENCES

Bennett, G. K., & Doppelt, J. E. Item difficulty and speed of response. Educational and Psychological Measurement, 1956, 16, 494-496.

Boag, A. K., & Neild, M. The influence of the time factor on the scores of the Triggs Diagnostic Reading Test as reflected in the performance of secondary school pupils as grouped according to ability. Journal of Educational Research, 1962, 55, 181-183.

Briggs, A., & Johnson, D. M. A note on the relation between persistence and achievement on the final examination. Journal of Educational Psychology, 1942, 33, 623-627.

Cronbach, L. J. Essentials of psychological testing. New York: Harper & Row, 1949.

Daly, J. L., & Stahmann, R. F. Effect of time limits on a university placement test. Journal of Educational Research, 1968, 62, 103-104.

Dubin, J. A., Osburn, H., & Winich, D. M. Speed and practice: Effects on Negro and white test performances. Journal of Applied Psychology, 1969, 53, 19-23.

Freeman, F. S. Power and speed: Their influence upon intelligence test scores. Journal of Applied Psychology, 1928, 12, 631-635.

Gates, A. I., & MacGinitie, W. H. Gates-MacGinitie reading tests. New York: Teachers College Press, 1965.

Glaser, R. Instructional technology and the measurement of learning outcomes: Some questions. American Psychologist, 1963, 18, 519-521.

Kendall, L. M. The effects of varying time limits on test validity. Educational and Psychological Measurement, 1964, 24, 789-800.

King, E. M., Lindquist, E. F., & Hieronymus, A. N. Canadian tests of basic skills. Toronto: Thomas Nelson, 1968.

Knapp, R. The effects of time limits on the intelligence test performance of Mexican and American subjects. Journal of Educational Psychology, 1960, 51, 14-20.

Longstaff, H. P., & Porter, J. P. Speed and accuracy as factors in objective tests in general psychology. Journal of Applied Psychology, 1928, 12, 636-642.

Michael, J. J., & Michael, W. B. The relationship of performance on objective achievement examinations to the order on which students complete them. Educational and Psychological Measurement, 1969, 29, 511-513.

Mollenkopf, W. G. An experimental study of the effects on item analysis data of changing item placement and test time limit. Psychometrika, 1950, 15, 291-315.

Mollenkopf, W. G. Time limits and the behavior of test takers. Educational and Psychological Measurement, 1960, 20, 223-230.

Stafford, R. E. The speededness quotient: A new descriptive statistic for tests. Journal of Educational Measurement, 1971, 8, 275-277.

Terranova, C. Relationship between test scores and test time. Journal of Experimental Education, 1972, 40, 81-83.

Thorndike, E. L. Fundamentals of learning. New York: Teachers College, 1932, 47. In Hilgard, E. R., & Bower, G. H., Theories of learning. New York: Appleton-Century-Crofts, 1966.

Wechsler, D. Wechsler intelligence scale for children. New York: Psychological Corporation, 1949.

Wechsler, D. Wechsler adult intelligence scale. New York: Psychological Corporation, 1955.

Wechsler, D. Wechsler intelligence scale for children -- revised. New York: Psychological Corporation, 1974.

Wesman, A. G. Some effects of speed in test use. Educational and Psychological Measurement, 1960, 20, 267-274.

Yates, A. J. The relationship between level and speed on two intelligence tests. British Journal of Educational Psychology, 1966, 36, 166-170, a.

Yates, A. J. Level, speed and personality factors in the intellectual performance of young children. British Journal of Educational Psychology, 1966, 33, 312-316, b.

APPENDIX A

EUROPE'S COMMON MARKET: PROBLEMS AND PROSPECTS

A common market is a joint venture by two or more nations designed to stimulate economic growth by eliminating trade tariffs. In 1957, France, Italy, West Germany, Belgium, the Netherlands and Luxembourg formed the Common Market. Three new members, Britain, Denmark and Ireland joined later. Since the total population is now 250 million the Common Market with its great internal market is now the strongest trading force on earth.

The Common Market provides trade free from tariffs between the members, cooperation in technological research and development, and free movement of both workers and money.

The British had to join because of economic problems at home. Some workers felt that the competition might cost them their jobs or raise food prices as high as those of the Germans or Italians. Norwegians votes against joining because they felt that their national interests may have to take second place to the Common Market interests. Each country that joined brought something. Britain brought industrial and scientific skills to the market. Denmark brought an agricultural industry that produces enough food for 15 million people.

France was the leader in the past but Britain is expected to play a major role now. Germany will probably hold the balance of power.

In spite of the fact that the Common Market is developing as an economic superpower each country is having internal problems. France's cities are becoming overcrowded, her rivers are being used as sewers which are polluting her coasts, and her workers complain of dull jobs.

Italy is hampered by crippling strikes. Britain is having civil strife and racial problems. All countries have inflation. The United States is becoming more competitive with the Common Market both in agricultural and non-agricultural products.

In spite of internal problems will the Common Market become a United States of Europe? This is the hope of Jean Monnet, a Frenchman, known as the father of the European community. Growing nationalism as well as a lack of a common defense policy hinder political union. Since decisions require unanimous consent, that is, consent of all members, it is likely that the Common Market will remain a loose federation of independent states.

APPENDIX B

EUROPE'S COMMON MARKET

1. The Common Market is a joint venture by two or more nations, designed to _____.
2. The six countries which first formed the Common Market were _____.
3. The enlarged Common Market added three members. These were _____.
4. The enlarged Common Market has a population equal to _____.
5. Together the nine Common Market countries make the _____ trading force on earth.
6. The greatest asset of the Common Market is its _____.
7. The asset on which the Common Market is built is similar to what other country? _____
8. The most efficient farmers are found in Denmark and _____.
9. Technological advances in the Common Market countries are expected to result from _____ in research and development.
10. Britain had to get involved in the Common Market because _____.
11. Some British workers felt that imports would offer competition to local products and may even threaten _____.

12. The British spend less on food than do Germans or _____

_____.

13. If a country is part of a larger community, national interests may have to take _____.

14. Which country was asked to join the Common Market in 1971 and refused? _____

15. Denmark produces enough food for _____ people.

16. Britain brought _____ and scientific skills to the Common Market.

17. For a long time _____ was considered the senior country of the Common Market.

18. Which of the new countries to join is expected to play a major role? _____

19. The balance of power is expected to be held by _____

_____.

20. The strongest industrial nation in the Common Market is _____

_____.

21. After World War II, Germany had to seek friendship especially with _____.

22. President Nixon has said that five economic superpowers: United States, Common Market, Soviet Union, China and _____ will determine the economic future of the world.

23. Each of the Common Market countries has internal problems. France has a problem with her cities because _____

_____.

24. France's coast is becoming polluted because _____

_____.

25. Some Frenchmen complain because office and factory jobs are _____ _____.

26. Italy's spectacular economic growth is hampered by _____ _____.

27. Britain is struggling with her own mini-Vietnam in _____ _____.

28. Besides civil strife Britain also has _____ _____ problems since she has received one million immigrants.

29. The rate of inflation in most European countries is _____ than in the United States.

30. Ludwig Erhard states that the feelings of unrest and uncertainty about all aspects of social life is largely due to the burning, destructive effects of growing _____ _____.

31. When wages are raised manufacturers in turn _____ _____ to absorb the cost of the higher wages they have to pay.

32. _____ is a Frenchman known as the father of the European community.

33. The Common Market has no defense policy. It relies on _____ _____.

34. Besides lack of a defense policy, the search for political unity is hindered by _____.

35. The French might be more receptive to political cooperation if _____.

36. The Common Market headquarters are located in _____ _____.

37. The Common Market is likely to remain mainly for _____ _____.

38. Any action by the Common Market requires _____ _____.

39. Instead of becoming an United States of Europe, the Common Market will probably remain a _____.

40. Some small Common Market countries are afraid of losing _____ _____ if the Common Market becomes a political state.