# THE EFFECTS OF TIME LIMITS AND READING COMPREHINSION ON PERFORMANCE ON THE CANADIAN TESTS OF BASIC SKILLS



FRED P. COLBERT

CANADIAN THESES ON MICROFICHE

I.S.B.N.

THESES CANADIENNES SUR MICROFICHE

National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada
Direction du développement des collections

Service des thèses canadiennes
sur microfiche

## 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.

## 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.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

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

NL-339 (r. 82/08)

Canada

# THE EFFECTS OF TIME LIMITS AND READING COMPREHENSION PERFORMANCE ON THE CANADIAN TESTS OF BASIC SKILLS

by

© Frederick Peter Colbert, B.A., 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

St. John's Newfoundland

ABSTRACT

The purpose of this study was to examine the effects of time limits and reading ability on performance on the Canadian Tests of Basic Skills (CTBS). To do this, the experimenter posed the following major questions: (1) Are the prescribed time limits for the CTBS adequate so that the results of these tests are valid indices of the abilities being tested? (2) Are differences in levels of reading comprehension associated with corresponding differences in levels of CTBS performance? (3) Is there an interaction between levels of reading ability and time limits which lead to misrepresentation of the abilities of specific groups of examinees?

Subjects were classified into groups of high, medium and low reading ability and were then randomly assigned to one of two possible time limit administrations of the CTBS. It was found that time conditions were responsible for no significant variation in performance of the subjects on the CTBS, except on the reading subtest. On all subtests, the extended time limit group surpassed the standard time limit group, but significant levels of variation were achieved only on the reading subtest.

It was also found that reading ability was a major factor in determining performance on the CTBS. The high reading ability group scored significantly higher than the medium and low groups on all subtests of the battery. The medium reading ability group, in turn, scored significantly higher than the low reading ability group, achieving significance on all subtests except punctuation and math computation.

On the entire battery, females outperformed males, achieving significant levels of variation on all subtests but reading and vocabulary.

On the basis of these findings, it was concluded that sex and reading ability were major determinants of performance on the CTBS for the subjects of this study. Time was not considered to be a major factor influencing performance except for the reading subtest.

## ACKNOWLEDGEMENTS

The writer wishes to express his grateful appreciation to the people who provided assistance and cooperation in conducting this study.

Gratitude is expressed to Dr. Bryan Hartmann for his advice, criticism and direction in the preparation of the study and the development of the report.

Thanks are expressed to Kay Colbert whose support and technical assistance were of profound importance.

My thanks are also extended to the Pentecostal School Board and to the principals and teachers of the schools involved who were so supportive of my work and this study.

Thanks to God whose presence and sustaining power made the entire effort possible.

Finally, I dedicate this thesis to my wife, Helen and to my sons, George and Michael, whose understanding and enthusiasm made the entire effort worthwhile.

TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

# CHAPTER I

# INTRODUCTION

## Purpose of the Study

The purpose of this study was to examine the effects of prescribed time limits and reading comprehension ability as factors determining performance on the Canadian Tests of Basic Skills (CTBS). The publishers of this test have not made clear just how the time limits prescribed for particular subtests were decided, nor have they provided an analysis of the effects of reading ability on performance on these tests.

Specifically, the researcher will try to answer the following questions:

1. Are the prescribed time limits for the CTBS adequate so that the results of these tests are valid indices of the abilities being tested?

2. Are differences in levels of reading comprehension associated with corresponding differences in levels of CTBS performance?

3. Is there an interaction between reading ability and time limits which may misrepresent the abilities of particular students?

Statement of the Problem

The researcher's involvement in this area grew out of an interest in the effects of prescribed time limits on the performance of students on various standardized tests. Research on the effects of time limits is neither exhaustive nor conclusive. Rather, the research done in this area seems to indicate that for some tests the prescribed limits are adequate for most examinees. The researcher could obviously conclude that some tests are sensitive to time limits and others are not. The problem could be solved if some formula existed for the calculation of time limits to be assigned to a particular test. Unfortunately, such a formula does not exist and the assignment of time limits to certain standardized test instruments may be a function of caprice and administrative convenience rather than an outcome of some systematic investigation of the effects of time limits. Indeed, it has been suggested that the popularity of test time limits may be due more to their practical administrative advantages than to any experimentally supported rationale governing the imposition of time limits on performance (Morrison, 1960).

Given that certain standardized tests may have inadequate time limits, the question arises as to why these time limits are inadequate. It would seem that in those cases where time limits are inadequate, the following conditions may prevail: some or all of the subjects do not have adequate time to attempt all of the items or they have time to attempt all of the items but do not have sufficient time to fully consider each one and thus make rash judgements.

One of the most obvious features of many standardized test instruments is the nature of their presentation to the subject. The test is generally given using the printed format, requiring at least a minimal level of reading ability. As such, the subject is confronted, in a sense, with two tests. This type of format requires the examinee to be not only knowledgeable in the domain of the test but also capable as a reader. In fact, it may be that performance on the test is significantly related to reading ability. If this is so, then the publisher would obviously be encumbered with a mandate to explain this relationship and its effect on test performance. If the potential relationship is ignored in cases where the test is as much a test of reading as anything else, then certainly the construct validity of such a test is jeopardized.

To add a further dimension to the discussion thus far, there may be a further effect on CTBS performance due to an interaction between various time limits and levels of reading comprehension; that is, these factors may operate in a unique way for distinct groups such that the effects of either variable will not be expressed in a systematic, predictable way for all groups of students. Consequently, test performance may not be interpretable except as part of an interaction between these two factors.

This issue may best be considered in terms of what it is about the nature of reading comprehension ability that makes it potentially sensitive to time limits. An explanation of the reading process will clarify the issue. Bormuth (1969) defined reading comprehension as a "set of generalized knowledge acquisition skills which permit people to acquire and exhibit information gained as a result of reading printed language" (p. 87). This definition stresses meaning interpretation skills.

Reading comprehension is the process of determining the meaning of written symbols. There are several skills that determine reading comprehension ability. Ekwall (1976) suggests that reading can best be understood in light of two constructs. He considers reading to be a product of the ability to recognize words and the ability to understand words and ideas. Ekwall claims that obtaining meaning from

words is considered to be the goal of reading. He further defines these two main constructs in this way: the ability to recognize words is composed of a number of sub-skills including the ability to recognize sight words; the ability to attack words and use decoding skills; the ability to use context and configuration clues; the ability to use phonetic analysis; and the ability to use structural analysis.

The ability to understand words and ideas is a product of vocabulary development and other comprehension skills. Ekwall includes under this heading Barrett's taxonomy of skills for comprehension. These skills are represented as abilities to extract meaning at the literal, inferential, evaluative and appreciative levels. From this discussion, it is evident that the reading process involves the use of various learned skills to establish meaning.

The skills involved in the reading process are all dependent on time. Applebee (1973) has suggested that there are at least four different skills which have a speed component. These skills may be classified as speeds of task performance, such as (i) speed of recognition of sight words, (ii) speed of decoding unknown words, (iii) speed of processing, and (iv) speed due to facility with structural analysis.

Indicating that speed should be considered as a dimension rather than a trait, Wesman (1960) points out that all activities are carried out over time and speed

should not be thought of as a fixed characteristic that a person applies to a task but, rather, as a factor whose value may change from time to time. It is readily apparent that the reading process is subject to this consideration and it may exhibit features consistent with this precept.

Yoakam (1955) pointed out that there are at least four distinct rates of reading, including: a skimming rate, a rapid reading rate, a work rate, and a careful reading rate. He suggests that, in general, a reader employs a particular rate in light of his purpose for reading which in turn is governed by a variety of factors, some of which are: ability, time available, motivation, and difficulty of the material. In summation, there are different work or reading rates characterized by the use of skills which themselves are subject to speed influences and, hence, time constraints. As well, a person's ability to reason and extract meaning at the literal, inferential, evaluative, or appreciative levels is a function of innate ability, i.e. ability to reason, and, like the other processes described, occurs over time.

Having identified these components of reading comprehension which are conducted over time, it is necessary to point out that because of this they are potentially sensitive to inadequate time limits; that is, where time is limited, students may not be able to use the skills

necessary to adequately comprehend a particular passage and their performance may be impaired.

Many studies in the past have looked at reading rate or reading speed and the relationship between time and comprehension (Husbands and Shores, 1950; Tinker, 1939). Their results consistently showed that rate and comprehension were highly correlated for easy material but the relationship was much less obvious as the difficulty of the material increased.

It seems that the relationship between rate and comprehension is not as fixed as first appears. Subsequent research has shown that whilst reading rate tends to be rigid rather than flexible, significant improvement in flexibility, without loss of accuracy, can be obtained when subjects are instructed to read faster (Harris and Sipay, 1980). Readers can also show some flexibility in modifying rates according to the kind of answer they are searching for and the difficulty of the material. Intermediate grade children tend to read faster for details than for main ideas and faster for main ideas than for order of events (Otto, Barrett and Harris, 1968). In 1975, Samuels and Dahl found significant differences in reading speed, with no loss in comprehension, when appropriate purposes were established for a group of fourth graders. This indicates that reading speed can be paced,

and fluctuation of the pace can occur without loss of comprehension. For students whose pace may be rather monotonous, performance on a given measure may be retarded when they are unaware that pace is an important goal of the exercise.

In summary, then, reading comprehension is a composite of various skills and abilities, all of which are exercised along a time dimension. The accuracy and speed with which one employs these skills seem to be mediated by the reader's purpose, the difficulty of the material, the reader's motivation, and his innate ability. Research has not shown exactly how these skills are related, but it has demonstrated that they certainly are mediating factors in the process of how rate affects comprehension. In the particular case of the CTBS with its prescribed time limits, there well may be some effect of the time limits on the reading comprehension of some or all students. Some students, whose rate of comprehension is slow, but whose power of comprehension is adequate, may exhibit unique variations in performance when time limits are imposed. The discovery of such an occurrence was one of the purposes of this study.

Significance

The impact of this study can best be understood in the light of the following factors, several of which are

germane to the Newfoundland context. Supporting evidence for statements one, two and four will be presented in the literature review below.

1. Research has demonstrated that the time limits of some standardized tests are inadequate such that they serve to impair the performance of some examinees.

2. Independent reports from several Newfoundland school districts indicate instances where over 50% of their students failed to complete certain CTBS subtests despite publishers' claims that the CTBS is a power test. This problem is most evident on the spelling, vocabulary, reading comprehension, map reading, capitalization and references subtests.

3. Newfoundland examinees' average CTBS performance is significantly below the Canadian average. Since the beginning of Newfoundland's CTBS testing program, the provincial composite score has not exceeded the thirty-first percentile in relation to the national norms. This result was achieved in 1981, and it indicates a gradual improvement in test performance over the last several years (Standards Testing Program Final Report, 1981).

4. It is suggested that Newfoundland examinees' low level of performance may be due to a general reading deficit in the student population (Seary, 1975).

This study may be significant for both the educational theorist and practitioner. It also may have implications of a theoretical and applied nature for those who are concerned with standardized testing and measurement.

At the theoretical level, the study offers further insight into an academic debate which is centered on the compromise which seems to have been reached in the old power versus speed controversy. This compromise seems to be a marriage of the basic assumptions of power and speed test theory. Such a union is based on the development of standardized instruments which follow a power test format and claim to be power tests, but which also include prescribed time limits. This study investigated the speededness of one of these measures, the CTBS. The results may assist the task of determining the extent to which such measures are in fact power tests rather than speeded tests.

At the most utilitarian level, the results of this study may have importance for those who rely on standardized test scores as bases for making educational decisions. Such decisions often relate to classroom placement, course choice, scholarship selection and occupational counselling. As well, the results of such tests often have important effects on teacher attitudes, self-conceptualization, as well as family and peer relationships.

In addition to these factors, the results of this study may be of importance to test constructors and publishers. Where power tests use time limits, the publisher should report the extent to which speed is a factor on the test (American Psychological Association Guidelines for Tests, 1974; Stafford, 1971; Myers, 1960). In scrutinizing the tests to be examined herein, little published information was available concerning why particular time limits were chosen and the extent to which they affect performance, if at all.

Furthermore, the instrument examined here, the CTBS, is currently the centrepiece of a province-wide testing program. As previously indicated, results thus far show that the mean scores of Newfoundland students are significantly below the Canadian norms on all subtests. This fact alone merits substantial investigation. As well, because reading may be a significant factor in explaining the performance of Newfoundland students on the CTBS, any information that can be provided by this investigation may be important.

## CHAPTER II

## REVIEW OF THE LITERATURE

The following review of the literature makes one immediately aware that the research conducted to date has not led to any consensus of opinion regarding the imposition of time limits on standardized tests. The time limits imposed on any instrument must be evaluated in light of the instrument in question. No conclusion can be generated regarding time limits in general.

This chapter utilizes the following format: (i) background to the problem; (ii) research studies involving time limits and race, rural/urban differences and disadvantaged groups, (iii) research studies involving time limits and measures of reading ability, and (iv) summary of literature review.

### Background to the Problem

The rationale for the present study has its roots in the early attempts to measure various aspects of human functioning. Some early efforts at measuring human ability were based on the assumption that the speed with which an individual performed a task was a measure of ability in

that domain; other efforts were based on the assumption that achievement was most accurately measured when no time limits were applied. Today, these are known as speed and power measures respectively. Recently it has been suggested that speed tests should be relegated to the evaluation of educational objectives where speed of response is an important objective; for example; typing and shorthand (Stanley and Hopkins, 1981).

Early research reported that speed and accuracy of response were interchangeable measures. Morrison (1960) summarized some conclusions reached by Spearman in 1927. On the basis of evidence gathered by himself and others, Spearman concluded that neither speed nor difficulty level contribute significant variance independent of 'g' (Bernstein, 1924; Hart and Spearman, 1914; Spearman, 1927). This theory might imply that it does not matter whether subjects' abilities are measured against a scale of difficulty or a scale or speed or some combination of the two. The true scores in the group should remain the same.

In addition to this, Hartmann (1980) reported that several early investigations reported positive correlations between speed and power test performances; therefore it was concluded that both types of measures yielded valid indices

of ability--usually mental ability (Freeman, 1928; Longstaff and Porter, 1928; McFarland, 1928, 1930; Ruch, 1924).

In contrast, other early investigations cited by Hartmann reported that the relationship between rates of response and mental ability was not significant (Bagley, 1901; Beck, 1932, 1933; Slater, 1933; Tinker, 1931, Tryon, 1933). In their opinion, speed was not a valid indicator of mental ability.

In more recent studies, this latter relationship has been substantiated by many researchers who report that there is a very low correlation between ability and working rates on tests (Tate, 1948; Ebel, 1954; Hopkins, 1964).

In addition to this, it has been further reported that scores obtained under speed conditions may differ in factor content from scores on the same test under power conditions (Baxter, 1941; Davidson and Carroll, 1945; Morrison, 1960), and that score variance on speed tests may contain separate portions attributable to the speededness of a test and its difficulty level. This finding is substantiated by Mollenkopf (1950, 1960) who indicated that a speed test measures speed and has little to do with measuring true ability. This is in contrast to the findings of Freeman et. al. (1928) who reported that either speed or accuracy measures were valid indices of ability.

Cronbach (1950) stated that speed is a legitimate element in achievement tests only when speed is an objective of the course. Speed is relevant and important in tests of typing, shorthand or tests of arithmetic for hiring cashiers. Speed is irrelevant if we wish to know how large a pupil's vocabulary is, how much science he knows or how accurately he can reason.

This latter reasoning seems to be inherent in the design of many recent test instruments. An attempt has been made to wed the underlying assumptions of a power test (provision of adequate time for all but the very weak to finish and item arrangement in ascending order of difficulty), with those of a speed test (provision of precise time limits in order to facilitate comparability of scores across different administrations). This arrangement appears to be evident in the CTBS, which is the subject of this investigation. The claim has been made by the editors of the CTBS that it is primarily a power test (King, 1975). This would suggest that the prescribed time limits would have little effect on performance.

One obvious aspect of the literature that was reviewed is that no general statement can be made regarding the use of time limits on standardized tests. It is readily apparent that where adequate time limits are observed, no significant differences will be found using

extended time. Rather, statements made in connection with time limits must be referenced to specific tests. This presents a problem for the researcher in that results obtained from different studies using different instruments are hardly comparable. It is of interest to note that a significant proportion of the studies covered in this review report significant score improvements under extended time conditions, suggesting that the problem of inadequate time limits may be widespread. However, no general statements may be made about time limits on all standardized tests.

Time Limits and Race, Rural/Urban Differences and Disadvantaged Groups

The studies reported in this section reveal no clear relationship between the imposition of time limits and racial origin. Knapp (1960) investigated the influence of time limits on the test scores of equal groups of Mexican and American subjects. Both groups were given the Cattell Culture Free Intelligence Test under power and speed conditions and the results were subjected to an analysis of variance. He found that both groups' scores were significantly higher under power conditions. In addition, he also determined that the higher-scoring Mexican group was more severely handicapped by the time limits than was the American group. Similar results were

reported in a study on tests of General Educational Development among adults in Texas conducted by Flores and Seaman in 1974. They reported that the overall performance of subjects was significantly higher under power conditions than under speed conditions. In particular, people under twenty, blacks and males achieved significantly higher on the power administration.

In direct contrast to their study was a study by Wild, Durso and Rubin (1982) which addressed the issue of test taking time and its effect on performance for special minority groups with the GRE national administration as a data base. They found that for neither the verbal nor the quantitative scores were average effects significantly different for subgroups defined by sex, race, or year out of school. They found less than a one point increase in performance in comparisons of standard time limit scores and extended time limit scores.

Between these extremes is a study by Dubin, Osburn and Winick (1969) who investigated the effects of extra test taking time on the performance of black subjects. The researchers hypothesized that black subjects would increase their performance relative to the performance of whites if time limits were extended and extra pretest practice was provided. In this study of 464 subjects, no significant racial difference was found but the performance of both

groups increased in relatively equal amounts under power conditions. Similar results have been reported by Reilly and Evans (1972). They administered the Quantitative Section of the Admission Test for Graduate Study in Business to groups of blacks and whites to determine if extending the time allotted would differentially affect the performance of black examinees. This hypothesis was rejected but the researchers did find a significant increase overall due to the main effect of timing.

From these studies, it was not apparent that race is a mediating factor in performance but four of the five studies cited did report a significant difference due to timing condition.

Race was not intended to be a major factor in this present study. However, it well may be that cultural factors are of significance in explaining Newfoundland's low performance on the CTBS. This study does not address the cultural issue directly as this factor is peripheral to the focus of the research. The studies investigating time limits and racial differences were included here in recognition of their valuable contribution to the question of time limits.

Kendall (1964) studied a group of 739 voluntary applicants for the Canadian Army using the Revised Examination M Test. His purpose was to determine the most adequate time limits for delivery of the above-mentioned instrument.

His findings indicated that unlimited time produced no better results than reduced time limits. He concluded that inadequate time limits impair validity but that unlimited time is wasted time, pointing to a need for some method of deriving time limits.

Miller and Weiss (1976) also investigated the question of rate and accuracy of performance under speed and power conditions. They found that when students were grouped by verbal ability, as defined by a vocabulary pretest, results indicated higher response rates under speed conditions. Verbal ability was assumed to be a good indicator of overall ability. However, they reported no significant difference for accuracy across time limit conditions, thusly indicating that the prescribed time limits were adequate. Like Kendall, they also concluded that unlimited test taking time produced no overall improvement in performance.

Similar results were reported by Barnett (1981). She studied the effects of the time limits of the Metropolitan Achievement Test on the performance of groups of learning disabled and emotionally disturbed children aged 10 to 12, with IQ above 80. She found that power conditions yielded slightly higher scores but the difference was not statistically significant.

Another special group was examined by Barnett (1981) with regard to determining the effect of time limits on test performance. In this study, a group of 306 deaf students were administered subtests of the Differential Aptitude Battery. A group of 170 completed all the items within the time limits; however, 136 subjects were identified who did not complete the items within the prescribed time limits. Thirty-six of these subjects were permitted to finish the test using extended time--the remainder were not. Results indicated no significant increase in mean scores for subjects who were permitted to finish relative to those who were not. There was no discussion as to why groups were divided in this way, nor was there any reason given as to why almost half of the total number tested did not finish within the prescribed time limits. Even though no differences were reported for the group of thirty-six which was permitted to finish the test, one wonders how the group was selected. In addition, one might question whether the group of 100 which was not permitted to finish would have exhibited similar performance had time limits been extended for them as well.

Moreton and Butcher (1964) studied urban and rural children in England and they investigated the hypothesis that rural children are at a disadvantage in the working of speeded tests. In the experiment, urban children

scored significantly higher than rural children overall. When subjects were given unlimited time, the difference decreased. The researchers concluded that speeded tests work to the advantage of urban children.

Similarly, Lewis (1973) studied a group of disadvantaged third grade students in Denver and varied test time limits and daily work limits to determine the effects on test performance. He found a group of slow responding, but accurate students whose performance on the Metropolitan Achievement Test increased significantly when time limits were extended.

This result is compatible with the findings of Yates who conducted two studies in 1966--the first with a group of first year engineering students, the second with a group of 100 children of nine years of age. In the first study he administered three tests to all subjects. These were: the Raven's Progressive Matrices, the Nufferno Level Test and an objective mathematics test. Results pointed to the existence of a slow working group which was handicapped by prescribed time limits. The significance of this study lay not in his discussion of extrovert-introvert characteristics of test takers but in the overall finding that the time limits can lead to differences in performance for special groups.

In the second study, Yates (1966) again discovered a group of slow but accurate workers whose intellectual level was severely underestimated by imposed time limits. His findings shed no light on reasons for this but his point was well made.

However, in apparent contrast to these findings is a study by Bridgeman (1980) who conducted several studies to investigate the differences between students labelled fast-accurate and slow-accurate. He found that on speeded measures, performance was not significantly different, suggesting that the prescribed time limits were adequate. His results indicated that for the measures used, the time limits were adequate and the slow-accurate students were able to work at a different tempo as the task required. These results supported findings by Bartis and Ford (1977) who found that children were able to modify their cognitive tempo in response to the demands of a task.

Studies of Time Limits and Reading Comprehension

As the purpose of the current study was to investigate the effects of time limits and levels of reading comprehension, the following studies were reviewed. In general, they present research which has demonstrated the effect of time limits on reading comprehension, and, in some cases, the effects of prescribed time limits on

measures of English expression.

Reilly and Evans have extensively researched the area of time limits and performance. A study by these authors in 1974 reported the effects of extended time limits on a test of reading comprehension. The subjects tested were from various ethnic backgrounds. Ethnicity was in fact the focus of the investigation. They used an unidentified national academic aptitude test and administered it under prescribed time limits of thirty minutes and extended time limits of forty minutes. The conclusions they drew were that reading comprehension performance increased overall but increases due to ethnic origin were not evident. In another study, in 1972, they administered a reading comprehension test to law school candidates of various socio-economic backgrounds. Again, the results pointed to a significant difference in performance when time limits were extended. This result is important for two reasons. It shows that some prescribed time limits are inadequate and that reading comprehension performance is subject to misrepresentation when inadequate time limits are imposed.

Forness and Dvorak (1982) reported that behaviourally disordered pupils are characterized by low achievement and variable test performance. In a study of forty adolescents with behavioural disorders tested under power and speed

conditions, performance increased only in the area of reading comprehension. This result served to point out again that reading comprehension scores can be hindered by inadequate time limits.

The results of the study by Forness and Dvorak are similar in many ways to findings reported by Wasson (1970). He administered the new Developmental Reading Tests to a sample of grade five pupils in a large suburban school district. The subtests were administered under standard and extended time conditions. An analysis of variance was employed and it was found that scores obtained under the extended time conditions significantly exceeded scores obtained under standard time conditions.

In a study of native American Indians, Immerman (1980) administered the Stanford Diagnostic Reading Test to two groups of applicants for admission to a polytechnical institute in New Mexico. The control group received the test under standard time conditions and the experimental group received it under untimed conditions. Significant improvement in test scores was reported for the experimental group, demonstrating that measures of reading ability are very sensitive to imposed time limits if these limits are inadequate.

A study by Boag and Neild (1962) in a New York high school was conducted by administering the Triggs Diagnostic

Reading Test to students in the upper grades. These students were grouped according to ability and were then randomly assigned to timed and untimed treatment conditions. The results showed an overall increase in performance between timed and untimed conditions marked by a significant interaction effect for students of average ability; that is, a certain group of students achieved significantly more during the untimed conditions. In effect, those students from the average group were hampered by the imposition of the prescribed time limits on this instrument. These findings were also supported by Yates (1966) and are similar to those reported in the following study.

Daly and Stahman (1968) reported on the admission test results for the University of Utah. The admission test was the Cooperative English Expression Test and was administered under prescribed time conditions. The study was conducted among a group of students who were assigned to the remedial English group as a result of the test. An alternative form of the test with extended time limits was given during the first class meetings. Results indicated that had the alternate form with extended time limits been applied, then 29 out of the 70 assigned to the remedial class would have received sufficiently high marks to avoid such placement. As the authors noted, these findings add validity to Yates' observation that time limits can be a hardship to many students.

Summary of Literature Review

This review of the literature has tended to focus on the wide variety of independent variables which sometimes interacts with time limits to determine test performance. The review has mainly tried to point out that in many instances the prescribed time limits may lead to the invalidation of test results. For groups of subjects whose ability is high but rate of work is slow the validity of test results on measures with prescribed time limits is particularly suspect (Yates, 1966; Boag and Neild, 1962; Daly and Stahman, 1968).

Studies using a wide variety of test instruments were reviewed. In many cases the use of additional test time served to produce higher achievement scores for the subjects. The fact that not all studies reported a difference due to time limits supported the previous contention that where adequate time limits are imposed, no time-related improvement in scores was likely to be found. This lack of concensus serves to highlight the need for test publishers to rigorously examine the way in which prescribed time limits are determined.

The review also pointed out that measures of reading comprehension may be particularly sensitive to time limits (Forness and Dvorak, 1982; Reilly and Evans, 1972, 1974; Boag and Neild, 1962). This was especially relevant for this study in that the examination of the nature of this potential relationship was one of the main purposes of the study.

# CHAPTER III

## PROCEDURE

### Introduction

This chapter utilizes the following format: (i) sample, (ii) experimental design, (iii) procedure, (iv) instruments, (v) hypotheses, (vi) limitations, (vii) summary.

### Summary

The sample for this study consisted of 165 students from six grade five classes from four schools in the Grand Falls area. These subjects were chosen because their schools were some of only a few not affected by an interruption of regular classes that occurred during the year. A labor dispute caused classes to be cancelled for a period of three weeks in almost all other schools in the province. Originally, the students from another school system had been selected for the study but due to the time they lost it was felt that their best interest would be served by selecting students that had not been so affected.

Subjects were classified as high, medium or low in reading ability according to the results of the Gates-MacGinitie

Reading Comprehension Subtest. Each group included a third of the entire distribution. For each reading level, subjects were randomly assigned to either the standard or extended time treatment of the CTBS.

Experimental Design

The design employed in this study was a modification of Campbell and Stanley's Post-Test-Only Control Group Design (Campbell and Stanley, 1969). It is represented in Figure 1.

| Reading Ability Group | Time Condition: Prescribed | Time Condition: Extended |
|---|---|---|
| Low | $X_1$ | $X_{23}$ |
| | $\vdots$ | $\vdots$ |
| | $X_{22}$ | $X_{44}$ |
| Medium | $X_{45}$ | $X_{67}$ |
| | $\vdots$ | $\vdots$ |
| | $X_{66}$ | $X_{88}$ |
| High | $X_{89}$ | $X_{123}$ |
| | $\vdots$ | $\vdots$ |
| | $X_{122}$ | $X_{144}$ |

Figure 1: Modification of Campbell and Stanley's Post-Test-Only Control Group Design

Procedure

All students were given the Gates-MacGinitie Reading Subtest several days prior to the administration of the CTBS. On the basis of this test, students were divided into high, medium and low reading ability groups. For each level, students were randomly assigned to either an extended time limit or standard time limit administration of the CTBS. The researcher was aided by an assistant in cases where it was necessary to divide classes by treatment groups. This was done to minimize the amount of time taken away from regular classroom instruction. However, the assistant used exactly the same standardized instructions as did the author and the researchers were randomly assigned when simultaneous testing was necessary.

The CTBS was administered in three sessions of two hours per session in each school. Two treatments occurred simultaneously in different classrooms. The standard time limit group was supervised by one researcher while the extended time limit group was supervised by the other. In some cases, the cooperating teacher agreed to supervise but was never involved in answering questions or giving instructions.

The standard time limit group referred to the group whose time for writing the test was exactly as prescribed by the test publishers. The extended time limit group referred to the group whose time for writing the test had

been extended by 25% over the prescribed time limits. This value was selected arbitrarily because research evidence on extended CTBS time limits was not available.

Instruments

The Canadian Tests of Basic Skills is a widely-used instrument in Newfoundland. It was developed from extensive work in test construction carried out at the University of Iowa over a thirty-five year period. Early in 1966 one of the U.S. editions was tried out in a special pilot project on a representative sample of Canadian schools. From this point onward, the Canadian Tests of Basic Skills has evolved into a nationally used and accepted instrument.

In this particular study, the Primary, Multi-Level Edition was used. Form 5 of Level 11 was the specific section employed. It consists of eleven subjects broken down into five main areas: (i) vocabulary, (ii) reading, (iii) language skills, (iv) work study skills, and (v) mathematic skills. In addition, a composite score was obtained by averaging these five areas. Each subtest consists of a series of multiple choice questions which require the subject to choose one answer and mark the corresponding space on a separate answer sheet.

The Canadian Edition of the Gates-MacGinitie Reading Tests is based upon the Second Edition of the Gates-MacGinitie Reading Tests published in the United States. These tests were published in Canada in 1980, and, even before Canadian

publication, were widely used in Canada. In this particular study, the Gates-MacGinitie Reading Comprehension Subtest was used. It consists of forty-three multiple choice items based upon sixteen different reading passages of varying lengths and difficulty levels. Level D, Form 1 was administered in this study.

Hypotheses

The following research hypotheses were used in this study:

1. There will be no significant difference between the means of the standard time limit group and the extended time limit group.

2. There will be no significant difference between the means of the three reading groups in terms of their CTBS performance.

3. There will be no significant interaction between the means of reading groups on standard time limit administrations and extended time limit administrations of the CTBS.

Limitations

This study dealt with only a small sample of a rather homogenous group of grade five students in four schools under

the jurisdiction of a single school board. The results, therefore, will only be generalizable to this Board at this grade level. In addition, the results will only apply to the CTBS and not to other timed tests.

Summary

This chapter consisted of a description of how the sample was chosen, how the independent groups were operationally defined and which instruments were employed. In addition, the chapter included a statement of hypotheses, a list of the procedures used, and a description of the experimental design involved.

# CHAPTER IV

## ANALYSIS OF RESULTS

### Introduction

The study was undertaken to determine if the time limits of the CTBS produced valid results for specific levels of reading ability as determined for the sample in question. In order to facilitate full treatment of the data available, a three-way analysis of variance was used, where time conditions, reading ability and sex were the independent variables involved. The level of significance was stated a priori at $p < .05$.

Because of losses due to mortality and incomplete data collection, the total number of subjects included in the analysis was reduced from 165 to 132. All statistical analyses were conducted using raw scores, grade equivalents and percentile ranks. All means and standard deviations quoted are expressed in percentile ranks in order to be consistent in the presentation of results for the various averaged scores. However, all findings reported were significant regardless of the type of score used and there was only one instance where a percentile rank was significant and its corresponding raw score was not. In this case the

result was not considered significant and the variation was attributed to a loss of accuracy due to calibration of the scales.

It should be noted in advance that the sample varies from the normal distribution. The scores are significantly below the Canadian norms on all subtests as demonstrated in Table 1. This low level of performance led the researcher to be extremely cautious in the interpretation of results and in the formulation of conclusions.

Hypothesis 3

Due to the nature of the findings, the third hypothesis will be considered first. This hypothesis stated that there would be no significant interaction among levels of reading ability and time limit conditions. In fact, this hypothesis could not be rejected as the findings revealed there was no significant interaction on any subtest for any combination of two factors or for the combination of all three factors. This finding made it possible for the researcher to focus the presentation of results on main effects. Table 2 is presented as a summation of the results of the three factor analysis of variance used in this study. It should be noted that this table was constructed from the data on the composite score. This score is an average of all the other subtests and

Table 1

Minimum, Maximum, Range, Mean, and Standard Deviation for Entire Sample by Subtest in Percentile Ranks

| Subtest | Minimum | Maximum | Range | Mean | Standard Deviation |
|---|---|---|---|---|---|
| Vocabulary | 1 | 89 | 88 | 23.9 | 22.4 |
| Reading | 1 | 91 | 90 | 26.9 | 24.8 |
| Spelling | 1 | 96 | 95 | 37.7 | 26.1 |
| Capitalization | 1 | 93 | 92 | 30.4 | 25.7 |
| Punctuation | 1 | 96 | 95 | 35.0 | 27.0 |
| Usage | 1 | 96 | 95 | 27.1 | 24.7 |
| Total Language | 1 | 96 | 95 | 30.5 | 26.7 |
| Visual Materials | 1 | 99 | 98 | 27.5 | 24.7 |
| References | 1 | 98 | 97 | 32.8 | 22.9 |
| Total Work | 1 | 99 | 98 | 28.8 | 25.2 |
| Math Concepts | 1 | 99 | 98 | 36.5 | 27.9 |
| Math Problems | 1 | 98 | 97 | 30.1 | 26.0 |
| Math Computation | 1 | 99 | 98 | 32.4 | 29.8 |
| Total Math | 1 | 96 | 95 | 32.8 | 28.0 |
| Composite | 1 | 98 | 97 | 26.3 | 10.1 |

Table 2

Analysis of Variance of Composite Results in Percentile Ranks

| Source | df | MS | F | Significance Level |
|---|---|---|---|---|
| Main Effects | | | | |
| Time | 1 | 792.44 | 2.90 | .09 |
| Reading Ability | 2 | 19156.93 | 70.17 | .00 |
| Sex | 1 | 2663.30 | 9.76 | .00 |
| 2-Way Interactions | | | | |
| Time by Reading Ability | 2 | 460.78 | 1.69 | .19 |
| Time by Sex | 1 | 3.49 | .01 | .91 |
| Reading Ability by Sex | 2 | 17.71 | .07 | .94 |
| 3-Way Interactions | | | | |
| Time by Reading Ability by Sex | 2 | 88.88 | .33 | .72 |
| Error | 120 | 273.01 | | |
| Total | 131 | 656.59 | | |

is considered to be a good indicator of overall performance on the CTBS. The findings presented are systematically representative of all the subtests. No significant interactions were found on any subtest. Similar data from the other subtests are included in Appendix A through Appendix M.

Hypothesis 1

The first hypothesis stated that there would be no significant differences in mean scores on the CTBS due to time conditions. This hypothesis was accepted. Nonsignificant results of all interaction terms involving the time factor permitted tests of the significance of the time main effects. With one exception these tests yielded nonsignificant results. Although mean scores on the extended time conditions exceeded mean scores on the standard time conditions on every subtest, significant differences were evident on only the reading subtest. Table 3 is a summary of means and standard deviations by subtest for both time conditions. Also, presented in Table 4 is a summary of the analysis of variance results for the reading subtest. All scores presented are in terms of percentile ranks.

Hypothesis 2

The second hypothesis stated that there would be no significant differences in performance on the CTBS due to differences in levels of reading ability. In every subtest, significant differences were found due to

Table 3

Means and Standard Deviations of Subtests by Time Condition in Percentile Ranks

| Subtest | Standard Time | | Extended Time | |
|---|---|---|---|---|
| | Mean | Standard Deviation | Mean | Standard Deviation |
| Vocabulary | 22.5 | 21.1 | 25.3 | 23.8 |
| Reading * | 23.0 | 23.0 | 30.8 | 26.2 |
| Spelling | 34.9 | 25.8 | 40.4 | 26.2 |
| Capitalization | 27.3 | 25.0 | 33.5 | 26.2 |
| Punctuation | 31.5 | 25.5 | 38.5 | 28.2 |
| Usage | 25.5 | 25.3 | 28.7 | 24.2 |
| Total Language | 27.2 | 25.6 | 33.9 | 27.6 |
| Visual Materials | 25.3 | 23.5 | 29.6 | 25.9 |
| References | 29.3 | 20.9 | 36.4 | 24.5 |
| Total Work | 25.4 | 22.8 | 32.2 | 27.2 |
| Math Concepts | 32.1 | 26.8 | 40.9 | 28.3 |
| Math Problems | 26.6 | 24.9 | 33.6 | 26.8 |
| Math Computation | 28.5 | 28.5 | 36.4 | 30.7 |
| Total Math | 28.0 | 26.0 | 37.7 | 29.3 |
| Composite | 22.6 | 23.5 | 30.0 | 27.3 |

* Subtests where significant differences due to time conditions were observed

Table 4

Analysis of Variance of Reading Results in Percentile Ranks

| Source | df | MS | F | Significance Level |
|---|---|---|---|---|
| Main Effects | | | | |
| Time | 1 | 1505.26 | 5.62 | .02 |
| Reading Ability | 2 | 19653.20 | 73.32 | .00 |
| Sex | 1 | 700.66 | 2.61 | .11 |
| 2-Way Interactions | | | | |
| Time by Reading Ability | 2 | 366.30 | 1.37 | .26 |
| Time by Sex | 1 | 236.79 | .88 | .35 |
| Reading Ability by Sex | 2 | 30.82 | .12 | .89 |
| 3-Way Interactions | | | | |
| Time by Reading Ability by Sex | 2 | 202.73 | .76 | .47 |
| Error | 120 | 268.04 | | |
| Total | 131 | 617.34 | | |

differences in levels of reading ability. Analysis of main effects for reading was possible due to the nonsignificant results of the interaction of reading ability with sex and time condition. Because every subtest was affected similarly, the researcher rejected the null hypothesis that no significant differences due to levels of reading ability would be found. Table 2, a summary of the analysis of variance for the composite score, represents this finding and it is pointed out that this finding is routinely present for all subtests. The analysis of variance results for the remaining subtests have been tabulated in Appendix A through Appendix M.

In this part of the analysis, Scheffé multiple comparisons were calculated to determine the significance of the differences among the three levels of reading ability. Due to the reputed conservative nature of the Scheffé procedure, a significance level of $p < .10$ was used in accordance with convention and Scheffé's (1959) recommendations.

The high reading ability subjects scored significantly higher than the medium and low groups on all subtests of the battery. The medium reading ability group, in turn, scored higher than the low reading ability group on all subtests except punctuation and math computation. Table 5 includes the means and standard deviations of subjects on each subtest for each level of reading ability. All cases where significant differences occurred between and among groups have been indicated.

Table 5

Means and Standard Deviations of Subtests by Levels of Reading Ability in Percentile Ranks

| Subtest | Low Reading Ability Group | | | Medium Reading Ability Group | | | High Reading Ability Group | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Standard Deviation | | Mean | Standard Deviation | | Mean | Standard Deviation |
| Vocabulary | 9.1 | 9.4 | * | 19.0 | 16.0 | * | 43.6 | 23.3 |
| Reading | 7.7 | 6.4 | * | 21.4 | 15.6 | * | 51.6 | 23.7 |
| Spelling | 21.3 | 18.7 | * | 36.1 | 21.5 | * | 55.6 | 25.6 |
| Capitalization | 14.7 | 17.0 | * | 26.4 | 19.0 | * | 50.1 | 26.4 |
| Punctuation | 21.6 | 20.5 | | 29.2 | 21.4 | * | 54.3 | 27.2 |
| Usage | 9.3 | 11.1 | * | 22.9 | 17.1 | * | 49.2 | 24.5 |
| Total Language | 12.3 | 15.2 | * | 25.4 | 18.7 | * | 53.9 | 26.1 |
| Visual Materials | 9.1 | 7.4 | * | 25.9 | 20.5 | * | 47.5 | 25.1 |
| References | 19.0 | 14.7 | * | 26.6 | 15.3 | * | 52.9 | 22.5 |
| Total Work | 10.1 | 10.6 | * | 24.5 | 18.1 | * | 51.8 | 24.0 |
| Math Concepts | 19.7 | 18.1 | * | 29.0 | 20.3 | * | 60.8 | 26.0 |
| Math Problems | 12.6 | 11.2 | * | 25.2 | 19.9 | * | 52.5 | 26.3 |
| Math Computation | 22.3 | 22.8 | | 21.2 | 20.7 | * | 53.8 | 32.4 |
| Total Math | 16.3 | 15.8 | * | 24.2 | 20.2 | * | 57.9 | 27.1 |
| Composite | 6.8 | 8.7 | * | 19.9 | 16.2 | * | 52.1 | 23.5 |

*Indicates significant differences between adjacent groups

This chapter also includes some a posteriori data analysis made possible by the method of data collection. The variable, sex, was included in the analysis and was found to be a significant factor in performance on the CTBS. This unexpected finding showed that girls outperformed boys on all subtests, with significant levels of difference found on all subtests except vocabulary and reading. The reader is again directed to Table 2 where a representative analysis of this result is presented. Table 6 presents the mean scores and standard deviations in percentile ranks for subjects on the subtests by sex.

Summary of Significant Findings

In summary, the major findings of the study were:

1. There was no interaction among time limit conditions, levels of reading ability and sex.
2. Time was not a significant factor in determining performance on the CTBS except in the case of the reading subtest.
3. Reading ability was a major factor in determining test performance on all subtests.
4. Females scored higher than males on all subtests, with statistically acceptable levels of significance

Table 6

Means and Standard Deviations of Subtests by Sex in Percentile Ranks

| Subtest | | Males Mean | Males Standard Deviation | Females Mean | Females Standard Deviation |
|---|---|---|---|---|---|
| Vocabulary | | 20.4 | 20.9 | 27.1 | 23.4 |
| Reading | | 20.7 | 23.0 | 32.6 | 25.3 |
| Spelling | * | 26.8 | 21.9 | 47.6 | 25.8 |
| Capitalization | * | 21.9 | 23.0 | 38.2 | 25.7 |
| Punctuation | * | 24.6 | 21.7 | 44.5 | 28.0 |
| Usage | * | 20.1 | 21.4 | 33.4 | 25.9 |
| Total Language | * | 19.7 | 21.3 | 40.5 | 27.5 |
| Visual Materials | * | 20.6 | 24.1 | 33.7 | 23.8 |
| References | * | 25.3 | 20.7 | 39.6 | 22.9 |
| Total Work | * | 20.4 | 23.9 | 36.4 | 24.1 |
| Math Concepts | * | 26.7 | 23.5 | 45.4 | 28.7 |
| Math Problems | * | 21.5 | 23.5 | 37.9 | 26.1 |
| Math Computation | * | 23.0 | 25.7 | 41.1 | 30.7 |
| Total Math | * | 21.7 | 23.1 | 42.9 | 28.4 |
| Composite | * | 18.0 | 22.2 | 33.8 | 26.4 |

*Subtests where significant differences due to sex were observed

being achieved on all but the vocabulary and reading subtests.

These findings have been expressed in Table 7, where a summary of significant main effects is presented.

Table 7

Summary of Levels of Significance for Main Effects by Factors for Subtests in Percentile Ranks

| Subtest | Time | Reading Ability | Sex |
|---|---|---|---|
| Vocabulary | .42 | .00 * | .57 |
| Reading | .02 * | .00 * | .10 |
| Spelling | .59 | .00 * | .00 * |
| Capitalization | .51 | .00 * | .00 * |
| Punctuation | .60 | .00 * | .00 * |
| Usage | .96 | .00 * | .00 * |
| Total Language | .48 | .00 * | .00 * |
| Visual Materials | .49 | .00 * | .02 * |
| References | .19 | .00 * | .00 * |
| Total Work | .22 | .00 * | .00 * |
| Math Concepts | .12 | .00 * | .00 * |
| Math Problems | .14 | .00 * | .01 * |
| Math Computation | .46 | .00 * | .00 * |
| Total Math | .10 | .00 * | .00 * |
| Composite | .09 | .00 * | .00 * |

* Subtests where significant differences due to main effects of variables were observed

# CHAPTER V

## SUMMARY, DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS

This chapter presents a discussion of the results in light of the major questions stated, the conclusions drawn from the results of the study, the implications for educators and some recommendations for future research.

### Summary

The purpose of the study was to examine the effects of time limits and reading ability as factors influencing performance on the Canadian Tests of Basic Skills. To do this, the experimenter posed these questions. Does extended time lead to higher mean scores for subjects? Does reading ability significantly affect performance on the CTBS? Do specific levels of reading ability interact with time limits to determine performance? To investigate these questions, the CTBS was administered under two time conditions to a sample of 165 grade five students from four schools. Subjects were tested using the Gates-MacGinitie Reading subtest, and, on the basis of this, were classified as high, medium or low reading ability subjects. From each reading level, subjects were then randomly assigned to

either a standard or extended time administration of the CTBS. The data were collected and analyzed using a three factor analysis of variance accounting for effects due to time conditions, reading ability, and sex. The results were reported in Chapter IV.

Discussion of Results

The analysis indicated that there were no significant differences due to time limit conditions on any subtest except reading, although on all subtests the extended group outperformed the standard time limit group. The fact that a significant difference was found on this subtest leads to at least two possible explanations. The first is Alpha error. This particular case is a special case of Alpha error that is experimentwise error. On the other hand, it may represent a real difference and would be consistent with the trend for extended time performance to be greater than standard time performance.

Forness and Dvorak (1982) reported in their study on low achievers that in the areas they tested under power and speed conditions that performance increased only for reading comprehension. In addition, Reilly and Evans' 1973 study pointed to a significant increase in reading comprehension scores when time limits were extended. Boag and Neild (1962) and Immerman (1980) also reported increases in reading scores when time limits were extended. This

would tend to support the conclusion that where reading ability is of primary importance on a test, time limits may lead to variations in performance. This finding suggests that on the reading comprehension subtest, the reading skills involved were sensitive to time and by extension, the reading skills required for the other subtests were not as sensitive to time. This result led the researcher to conclude that time limits, as prescribed by the publishers, are adequate for the majority of the subtests. With regard to the observation noted above that over 50% of Newfoundland students fail to complete certain CTBS subtests, this conclusion indicates that for these students all subtests except reading should be interpreted in accord with the publisher's claim that the measure is a power test.

The results of the Scheffé comparisons indicated that the high reading ability subjects scored significantly higher than the medium and low reading groups on all subtests of the battery. The medium reading ability group, in turn, scored higher than the low reading ability group on all subtests of the battery, achieving statistically acceptable levels of significance on all subtests, except punctuation and math computation. This would seem to be a reasonable finding in light of the nature of the subtest involved where reading skills are not meant to be the focus of the subtest. This lack of significance between low and medium reading ability

groups on the subtests is not thought to impinge on the generality of this result.

The fact that reading ability is such a significant component of performance on the CTBS is noteworthy. It is not unreasonable to explain such an effect, however, when it is considered that much of the information children obtain in school is gained through the printed medium. Hence, better readers typically learn more and can therefore perform at higher levels. In addition, the influence of this reading factor may be better understood in a study where intellectual level is accounted for. This finding indicates a need for further investigation of this factor. The significant reading ability effect questions the construct validity of several of the CTBS subtests. Until the relationship between reading ability and performance on the CTBS subtests is clarified, it is difficult to interpret the results of individual subtests of basic skills. At this point it is sufficient to note that this result is very compatible with Seary's (1975) hypothesis that the generally poor CTBS performance of Newfoundland students may be due to a general reading deficit in the student population.

The analysis also indicated that girls significantly outscored boys on all subtests of the battery except two, Reading and Vocabulary. This factor had not been anticipated a priori, but subsequent investigation of the relevant literature revealed that a majority of studies have shown

that girls tend to achieve higher results than boys at the elementary level (Kolesnik, 1970). This difference may be accounted for in a number of ways, including: neurological, environmental and sex role factors. Furthermore, according to Waetjen (1962) and Maccoby (1966), boys outnumber girls as underachievers by two or three to one. In light of the fact that this sample scored well below the Canadian norms, it might be concluded that it is a unique sample of elementary students characterized by a fairly large number of under-achieving males. Obviously, further investigation of the sex factor is necessary in order to more fully understand its influence on performance with groups more representative of the normal distribution.

The potential interaction that was posited in the third hypothesis, that certain reading groups might be more influenced by inadequate time limits than others, was not supported in this study. The results do not support the findings of Boag and Neild (1962), Yates (1966), Daly and Stahman (1968) or Immerman (1980) who found that certain subgroups were more prone than other groups to be handicapped by inadequate time limits. It should be noted that these other studies concerned themselves with other variables and test instruments than those investigated here and that the findings of this study are not directly comparable with the results reported by these other authors.

## Conclusions

The conclusions reached in this study are limited to the CTBS performance of grade five students within four schools of a school board located in central Newfoundland. Because of the apparently unique sample used in the study, generalization beyond the above limits is especially inappropriate.

The conclusions of this investigation are as follows:

1. Prescribed time limits are not a determining factor of performance on most CTBS subtests. On tests of reading comprehension, or on subtests which focus sharply on reading skills, time limits may be a mediating factor determining performance.

2. Reading ability is a significant factor on all CTBS subtests. Reading ability is directly related to test performance. This may be intrinsically linked with intellectual level, and further investigation is required in order to fully understand this relationship.

3. Girls tend to score significantly higher than do boys on all but the reading and vocabulary subtests of the CTBS. This finding was unexpected and could not be anticipated by inspection of the test manuals or supplements in advance of the study.

## Recommendations

As a result of the study, the following recommendations are made for further research and application:

1. A similar study should be conducted with a sample more characteristic of the normal distribution. In a future study, intelligence as a factor should be incorporated into the design. A further variation on this would be to control sex in such a way as to clarify the role of sex as a determinant of performance.

2. Similar studies should be conducted to determine whether the time limits are adequate on other widely-used standardized tests.

3. Teachers should be aware that arbitrarily assigned time limits may not lead to maximum success for all students on teacher-made tests.

4. Test users may find it beneficial to consider the possible effects of reading ability, time limits and sex on test performance when attempting to interpret performance results.

# REFERENCE NOTES

## REFERENCE NOTES

Hartmann, B. Test time limits and performance on group tests of intelligence and educational achievement. Unpublished manuscript, Memorial University of Newfoundland, 1980.

Fagan, M. Standards Testing Program Final Report 1981. Newfoundland: Government of Newfoundland and Labrador, Department of Education, Testing Branch, 1982.

# REFERENCE LIST

## REFERENCE LIST


American Psychological Association Guidelines for Tests. In O.K. Buros (Ed.), Tests in print II. New Jersey: The Gryphon Press, 1974.

Applebee, R.R. High school English instruction today. New York: Appleton, Century, Crofts, 1973.

Bagley, W.C. "On the correlation of mental and motor ability in school children." American Journal of Psychology XII (1901), 193-205.

Barnett, E. Performance of inner city learning disabled and emotionally disturbed youth on power and timed achievement tests. (University of Connecticut, 1981). Dissertation Abstracts International, 1981, 41, 4522A-5271A.

Bartis, S.W., & Ford, L.H. Reflective - impulsivity, conservation and the development of central cognitive tempo. Child Development, 1977, 48, 953-959.

Baxter, B.M. An experiment of the contribution of speed and level in an intelligence test. Journal of Educational Psychology, 1941, 32, 285-296.

Beck, L.F. Relation of speed of reaction to intelligence. American Journal of Psychology, 1932, 44, 793-795.

Beck, L.F. The role of speed in intelligence. Psychological Bulletin, 1933, 30, 169-178.

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.

Bormuth, J.R. (Operational definition of comprehensive instruction). In K.S. Goodman & J.T. Fleming (Eds.), Psycholinguistics and teaching of reading. Delaware: International Reading Association, 1969.

Bridgeman, B. Generality of fast or slow test taking style across a variety of cognitive tasks. Journal of Educational Measurement, 1980, 17 (3).

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.

Campbell, D.T., & Stanley, J.C. Experimental and quasi-experimental design for research. Boston: Houghton-Mifflin Co., 1963.

Cronbach, L.J. Response sets and test validity. Educational and Psychological Measurement, 1950, 10, 3-31.

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

Davidson, W.M., & Carroll, J.B. Speed and level components in time limit scores, a factor analysis. Educational and Psychological Measurement, 1945, 5, 411-427.

Dubin, J.A., Osburn, H., & Winick, D.A. Speed and practice: Effects of negro and white test performance. Journal of Applied Psychology, 1969, 53, 19-23.

Ebel, R.L. Procedures for the analysis of classroom tests. Educational and Psychological Measurement, 1954, 14, 352-364.

Ekwall, E.E. Diagnosis and remediation of the disabled reader. Boston: Allyn and Bacon, Inc., 1976.

Flores, T., & Seaman, D. A comparative study of adult student performance on timed versus non-timed GED tests in Texas. (ERIC Document Reproduction Service No. ED 154 138, 1974).

Forness, S., & Dvorak, R. Effects of test time-limits on achievement scores of behaviourally disordered adolescents. Behavioural Disorders, 1982, 17 (4), 207-212.

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

Glass, G.V., & Stanley, J.F. Statistical methods in education and psychology. New Jersey: Prentice Hall, Inc., 1970.

Gulliksen, H. Speed versus power tests. Theory of Mental Tests. New York: Wiley, 1950.

Harris, A.J., & Sipay, E.R. How to increase reading ability. New York : Longman, Inc., 1980.

Hopkins, K. Extrinsic reliability: Estimating and attenuating variance from response styles, chance and other irrelevant sources. Educational and Psychological Measurement, 1964, 24, 271-281.

Husbands, K.L., & Shores, J.F. Are fast readers the best readers? Elementary English, 1950, 27, 52-57.

Immerman, M.A. The effects of eliminating time restraints on a standardized test with American Indian adults. (ERIC Document Reproduction Service No. ED 196 584, 1980).

Kahn, S.B. The relative magnitude of speed and power in SCAT. Journal of Educational Measurement, 1968, 5, 327-330.

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

King, E. Canadian tests of basic skills. Canada: Thomas Nelson & Sons, 1975.

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

Kolesnik, W.B. Educational psychology. New York: McGraw Hill, Inc., 1970.

Lewis, W.E. Effects of rate and accuracy of test responses, removal of test time-limits and teacher expectancies on achievement test scores of disadvantaged third grade students in Denver. (Oregon State University, 1973). Dissertation Abstracts International, 1973, 34, 2073A-2808A.

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.

Maccoby, E.E. The development of sex differences. California: Stanford University Press, 1966.

MacGinitie, W.H. Gates-MacGinitie Reading Tests. Canada: Nelson Canada Ltd., 1980.

McFarland, R.A. The role of speed in mental ability. Psychological Bulletin, 1928, 30, 595-612.

McFarland, R.A. An experimental study of the relationship between speed and mental ability. Journal of General Psychology, 1930, 3, 67-97.

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.

Miller, T., & Weiss, D. Effect of time limits on test taking behaviour, 1976. (ERIC Document Reproduction Service No., ED 126 123).

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

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

Moreton, C.A., & Butcher, H.J. Are rural children handicapped by the use of speeded tests in selection procedures? British Journal of Educational Psychology, 1963-64, 33-34, 22-30.

Morrison, E.J. On test variance and the dimensions of the measurement situation. Educational and Psychological Measurement, 1960, 2, 231-251.

Myers, C.T. Symposium: The effects of time limits on test scores. Educational and Psychological Measurement, 1960, 20, 221-222.

Otto, W., Barrett, T.C., & Harris, T.L. Research in reading. Journal of Experimental Education, 1968, 33, 65-67.

Reilly, R.R., & Evans, F.R. A study of speededness as a source of test biases. Journal of Educational Measurement, 1972, 9, 123-131.

Reilly, R.R., & Evans, F.R. A study of test speededness as a potential source of bias in the quantitative score of the Admissions Test for Graduate Study in Business. Research in Higher Education, 1973, 1, 173-183.

Reilly, R.R., & Evans, F.R. The effects of test time limits on performance of culturally defined groups. (ERIC Document Reproduction Service No. ED 102 276, 1974).

Ruch, G.M. The speed factor in mental measurements. Journal of Educational Research, 1924, 9, 39-45.

Samuels, S.J., & Dahl, P.R. Establishing appropriate purpose for reading and its effect on flexibility on reading rate. Journal of Educational Psychology, 1975, 67 (1), 38-43.

Scheffé, H. The analysis of variance. New York: John Wiley, 1959.

Seary, A.A. Standards testing program. Government of Newfoundland and Labrador, Department of Education, St. John's, Newfoundland, 1975.

Slater, P. Speed of work in intelligence tests. British Journal of Psychology, 1933, 29, 55-68.

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

Stanley, J.F., & Hopkins, K. Educational and psychological measurement and evaluation. New Jersey: Prentice Hall, Inc., 1981.

Tate, M.W. Individual differences in speed of response in mental test materials of varying degrees of difficulty. Educational and Psychological Measurement, 1948, 8, 353-374.

Tate, M.W. Notes on the measurement of mental speed. Journal of Educational Psychology, 1950, 41, 219-236.

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

Tinker, M.A. The significance of speed in test response. Psychological Review, 1931, 38, 450-454.

Tinker, M.A. Speed versus comprehension in reading as affected by level of difficulty. Journal of Educational Psychology, 1939, 30, 81-94.

Tinker, M.A. Rate of work in reading performance as measured in standardized tests. Journal of Educational Psychology, 1945, 36, 217-228.

Tinker, M.A., & McCullough, C.M. Teaching elementary reading. New York: Meredith Corporation, 1966.

Tryon, C.M. On the nature of speed and its relation to other variables. Journal of General Psychology, 1933, 8, 192-216.

Waetjan, W.B. Is learning sexless? NEA Journal, 1962, 51, 12-14.

Wasson, J.B. The influence of time limits on reading comprehension tests among fifth grade pupils of differing abilities. (University of Minnesota, 1969). Dissertation Abstracts International, 1970, 30, 3577A-4602A.

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

Wild, C.L., Durso, R., & Rubin, D.B. Effect of increased test-taking time on test scores by ethnic group, years out of school and sex. Journal of Educational Measurement, 1982, 19 (1).

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.

Yoakam, G.A. Basal reading instruction. New York: McGraw Hill Book Co., Inc., 1955.

# APPENDICES

## Appendix A

### Analysis of Variance of Vocabulary Results in Percentile Ranks

| Source | df | MS | F | Significance Level |
|---|---|---|---|---|
| Main Effects | | | | |
| Time | 1 | 191.86 | .67 | .42 |
| Reading Ability | 2 | 12333.90 | 42.88 | .00 |
| Sex | 1 | 91.87 | .32 | .57 |
| 2-Way Interactions | | | | |
| Time by Reading Ability | 2 | 866.82 | 3.01 | .05 |
| Time by Sex | 1 | 630.20 | 2.20 | .14 |
| Reading Ability by Sex | 2 | 273.75 | .95 | .39 |
| 3-Way Interactions | | | | |
| Time by Reading Ability by Sex | 2 | 596.15 | 2.07 | .13 |
| Error | 120 | 287.67 | | |
| Total | 131 | 502.86 | | |

## Appendix B

### Analysis of Variance of Spelling Results in Percentile Ranks

| Source | df | MS | F | Significance Level |
|---|---|---|---|---|
| Main Effects | | | | |
| Time | 1 | 129.07 | .30 | .59 |
| Reading Ability | 2 | 8965.73 | 20.83 | .00 |
| Sex | 1 | 8398.55 | 19.51 | .00 |
| 2-Way Interactions | | | | |
| Time by Reading Ability | 2 | 601.63 | 1.40 | .25 |
| Time by Sex | 1 | 276.97 | .64 | .42 |
| Reading Ability by Sex | 2 | 177.85 | .41 | .66 |
| 3-Way Interactions | | | | |
| Time by Reading Ability by Sex | 2 | 101.57 | .24 | .79 |
| Error | 120 | 430.49 | | |
| Total | 131 | 679.42 | | |

## Appendix C

### Analysis of Variance of Capitalization Results in Percentile Ranks

| Source | df | MS | F | Significance Level |
|---|---|---|---|---|
| Main Effects | | | | |
| Time | 1 | 182.93 | .44 | .51 |
| Reading Ability | 2 | 9841.74 | 23.79 | .00 |
| Sex | 1 | 4311.89 | 10.42 | .00 |
| 2-Way Interactions | | | | |
| Time by Reading Ability | 2 | 352.92 | .85 | .43 |
| Time by Sex | 1 | 83.27 | .20 | .65 |
| Reading Ability by Sex | 2 | 525.73 | 1.27 | .28 |
| 3-Way Interactions | | | | |
| Time by Reading Ability by Sex | 2 | 905.88 | 2.19 | .12 |
| Error | 120 | 413.66 | | |
| Total | 131 | 661.30 | | |

## Appendix D

### Analysis of Variance of Punctuation Results in Percentile Ranks

| Source | df | MS | F | Significance Level |
|---|---|---|---|---|
| Main Effects | | | | |
| Time | 1 | 127.12 | .27 | .60 |
| Reading Ability | 2 | 7993.14 | 17.11 | .00 |
| Sex | 1 | 8064.66 | 17.26 | .00 |
| 2-Way Interactions | | | | |
| Time by Reading Ability | 2 | 720.47 | 1.54 | .22 |
| Time by Sex | 1 | 842.38 | 1.80 | .18 |
| Reading Ability by Sex | 2 | 412.03 | .88 | .42 |
| 3-Way Interactions | | | | |
| Time by Reading Ability by Sex | 2 | 1199.84 | 2.57 | .08 |
| Error | 120 | 467.23 | | |
| Total | 131 | 729.06 | | |

## Appendix E

### Analysis of Variance of Usage Results in Percentile Ranks

| Source | df | MS | F | Significance Level |
|---|---|---|---|---|
| Main Effects | | | | |
| Time | 1 | 1.02 | .00 | .96 |
| Reading Ability | 2 | 12847.16 | 39.35 | .00 |
| Sex | 1 | 2439.04 | 7.47 | .01 |
| 2-Way Interactions | | | | |
| Time by Reading Ability | 2 | 415.32 | 1.27 | .28 |
| Time by Sex | 1 | 49.36 | .15 | .70 |
| Reading Ability by Sex | 2 | 284.91 | .87 | .42 |
| 3-Way Interactions | | | | |
| Time by Reading Ability by Sex | 2 | 584.02 | 1.79 | .17 |
| Error | 120 | 326.53 | | |
| Total | 131 | 609.69 | | |

## Appendix F

### Analysis of Variance of Total Language Results in Percentile Ranks

| Source | df | MS | F | Significance Level |
|---|---|---|---|---|
| Main Effects | | | | |
| Time | 1 | 171.14 | .48 | .49 |
| Reading Ability | 2 | 13444.22 | 37.92 | .00 |
| Sex | 1 | 7696.58 | 21.71 | .00 |
| 2-Way Interactions | | | | |
| Time by Reading Ability | 2 | 693.60 | 1.96 | .15 |
| Time by Sex | 1 | 225.13 | .64 | .43 |
| Reading Ability by Sex | 2 | 136.86 | .39 | .68 |
| 3-Way Interactions | | | | |
| Time by Reading Ability by Sex | 2 | 644.05 | 1.82 | .17 |
| Error | 120 | 354.53 | | |
| Total | 131 | 714.76 | | |

## Appendix G

### Analysis of Variance of Visual Materials Results in Percentile Ranks

| Source | df | MS | F | Significance Level |
|---|---|---|---|---|
| Main Effects | | | | |
| Time | 1 | 178.09 | .49 | .49 |
| Reading Ability | 2 | 13007.88 | 35.51 | .00 |
| Sex | 1 | 1925.87 | 5.26 | .02 |
| 2-Way Interactions | | | | |
| Time by Reading Ability | 2 | 272.35 | .74 | .48 |
| Time by Sex | 1 | 5.91 | .02 | .90 |
| Reading Ability by Sex | 2 | 182.14 | .50 | .61 |
| 3-Way Interactions | | | | |
| Time by Reading Ability by Sex | 2 | 125.39 | .34 | .71 |
| Error | 120 | 366.27 | | |
| Total | 131 | 612.09 | | |

## Appendix H

### Analysis of Variance of References Results in Percentile Ranks

| Source | df | MS | F | Significance Level |
|---|---|---|---|---|
| Main Effects | | | | |
| Time | 1 | 476.16 | 1.69 | .20 |
| Reading Ability | 2 | 10280.25 | 36.49 | .00 |
| Sex | 1 | 2914.73 | 10.35 | .00 |
| 2-Way Interactions | | | | |
| Time by Reading Ability | 2 | 810.09 | 2.88 | .06 |
| Time by Sex | 1 | 8.21 | .03 | .87 |
| Reading Ability by Sex | 2 | 530.42 | 1.88 | .16 |
| 3-Way Interactions | | | | |
| Time by Reading Ability by Sex | 2 | 485.13 | 1.72 | .18 |
| Error | 120 | 281.71 | | |
| Total | 131 | 526.07 | | |

## Appendix I

### Analysis of Variance of Total Work Results in Percentile Ranks

| Source | df | MS | F | Significance Level |
|---|---|---|---|---|
| Main Effects | | | | |
| Time | 1 | 470.58 | 1.52 | .22 |
| Reading Ability | 2 | 15105.19 | 48.80 | .00 |
| Sex | 1 | 3301.72 | 10.67 | .00 |
| 2-Way Interactions | | | | |
| Time by Reading Ability | 2 | 779.97 | 2.52 | .09 |
| Time by Sex | 1 | 4.00 | .01 | .91 |
| Reading Ability by Sex | 2 | 150.24 | .49 | .62 |
| 3-Way Interactions | | | | |
| Time by Reading Ability by Sex | 2 | 247.60 | .80 | .45 |
| Error | 120 | 309.54 | | |
| Total | 131 | 635.49 | | |

## Appendix J

### Analysis of Variance of Math Concepts Results in Percentile Ranks

| Source | df | MS | F | Significance Level |
|---|---|---|---|---|
| **Main Effects** | | | | |
| Time | 1 | 1046.00 | 2.43 | .12 |
| Reading Ability | 2 | 15232.69 | 35.33 | .00 |
| Sex | 1 | 4928.22 | 11.43 | .00 |
| **2-Way Interactions** | | | | |
| Time by Reading Ability | 2 | 430.91 | 1.00 | .37 |
| Time by Sex | 1 | 66.94 | .16 | .69 |
| Reading Ability by Sex | 2 | 586.76 | 1.36 | .26 |
| **3-Way Interactions** | | | | |
| Time by Reading Ability by Sex | 2 | 41.44 | .10 | .91 |
| Error | 120 | 431.11 | | |
| Total | 131 | 775.76 | | |

## Appendix K

### Analysis of Variance of Math Problems Results in Percentile Ranks

| Source | df | MS | F | Significance Level |
|---|---|---|---|---|
| Main Effects | | | | |
| Time | 1 | 863.43 | 2.23 | .14 |
| Reading Ability | 2 | 14839.65 | 38.24 | .00 |
| Sex | 1 | 3249.85 | 8.38 | .01 |
| 2-Way Interactions | | | | |
| Time by Reading Ability | 2 | 1.87 | .01 | 1.00 |
| Time by Sex | 1 | 10.41 | .03 | .87 |
| Reading Ability by Sex | 2 | 258.93 | .67 | .52 |
| 3-Way Interactions | | | | |
| Time by Reading Ability by Sex | 2 | 46.81 | .12 | .89 |
| Error | 120 | 388.04 | | |
| Total | 131 | 678.05 | | |

## Appendix L

### Analysis of Variance of Math Computation Results in Percentile Ranks

| Source | df | MS | F | Significance Level |
|---|---|---|---|---|
| **Main Effects** | | | | |
| Time | 1 | 330.83 | .54 | .46 |
| Reading Ability | 2 | 9806.18 | 16.05 | .00 |
| Sex | 1 | 6463.27 | 10.58 | .00 |
| **2-Way Interactions** | | | | |
| Time by Reading Ability | 2 | 180.74 | .30 | .75 |
| Time by Sex | 1 | 50.30 | .08 | .78 |
| Reading Ability by Sex | 2 | 731.79 | 1.20 | .31 |
| **3-Way Interactions** | | | | |
| Time by Reading Ability by Sex | | 1156.05 | 1.89 | .16 |
| Error | 120 | 611.14 | | |
| Total | 131 | 885.65 | | |

Appendix M

Analysis of Variance of Total Math Results in Percentile Ranks

| Source | df | MS | F | Significance Level |
|---|---|---|---|---|
| Main Effects | | | | |
| Time | 1 | 1106.06 | 2.72 | .10 |
| Reading Ability | 2 | 15649.47 | 38.52 | .00 |
| Sex | 1 | 7086.05 | 17.44 | .00 |
| 2-Way Interactions | | | | |
| Time by Reading Ability | 2 | 12.20 | .03 | .97 |
| Time by Sex | 1 | 82.26 | .20 | .65 |
| Reading Ability by Sex | 2 | 441.38 | 1.09 | .34 |
| 3-Way Interactions | | | | |
| Time by Reading Ability by Sex | 2 | 109.22 | .27 | .77 |
| Error | 120 | 406.25 | | |
| Total | 131 | 785.58 | | |