Loyola University Chicago
Loyola eCommons

Dissertations | Theses and Dissertations

1967

A Study of Deliberate Faking in the MMPI with Seminarians

John Gerald Grant
Loyola University Chicago

Follow this and additional works at: https://ecommons.luc.edu/luc_diss

 Part of the Psychology Commons

Recommended Citation

Grant, John Gerald, "A Study of Deliberate Faking in the MMPI with Seminarians " (1967). *Dissertations*. 833.
https://ecommons.luc.edu/luc_diss/833

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Dissertations by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License.
Copyright © 1967 John Gerald Grant

AN INVESTIGATION INTO THE EFFECTS OF IMMEDIATE KNOWLEDGE OF PERFORMANCE AND SUBSEQUENT TEST RESULTS

by

Thomas John Ginley

A Dissertation Submitted to the Faculty of the Graduate School of Loyola University in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

February

1967

LIFE

Thomas John Ginley was born in Chicago, Illinois, on January 25, 1938.

He was graduated from Loyola Academy in June, 1956 and Loyola University, Chicago, Illinois, June, 1960 with the degree of Bachelor of Science. He began his graduate studies in the Department of Psychology at Loyola University in June, 1960 and earned a Master of Arts degree in June, 1963.

The author was employed as a member of the test construction division of the Chicago Civil Service Commission, December, 1960 to February, 1963. He is presently employed as Assistant Secretary, Council on Dental Education, and Director, Division of Educational Measurements, American Dental Association.

ACKNOWLEDGEMENTS

The writer wishes to express his gratitude to Dr. Edmund P. Marx, Professor in the Department of Psychology, Loyola University for his encouragement and advice in the preparation of this study.

To Dr. William P. Schoen, Jr., Dean, Loyola University School of Dentistry, Dr. George W. Teuscher, Dean, Northwestern University Dental School, Dr. Seymour H. Yale, Dean, University of Illinois College of Dentistry, Dr. William R. Mann, Dean, University of Michigan School of Dentistry and Dr. Leroy R. Boling, Dean, Washington University School of Dentistry, are due most sincere thanks for securing subjects for this research project.

The author also wishes to thank Mrs. Gene Childress and Mr. Robert Pyskacek for their assistance in the preparation of this study as well as special thanks to Mrs. Harriet Tuchten for the effort employed in carefully typing this manuscript.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

PROBLEM

The developments pertaining to the measurement of man's abilities have their foundations in antiquity. The ordering of man on the basis of an ability measure has been attempted throughout the course of history. DuBois (1964) noted that in 2200 B.C., the emperor of China is said to have examined his government officials periodically and either promoted or dismissed them from service on the basis of the results. The degree of sophistication related to the method employed in measurement demonstrated a spiraling effect as new methods and procedures were devised. History alone, however, cannot provide the sufficient maturing influence necessary for an adequate measurement system within the field of psychological testing. Both the refinement of old and the development of new concepts are essential mandates for any satisfactory growth within the field of psychometrics.

The basic function of psychological testing is to measure differences between individuals or between performances of the same individual under different conditions. The initial appli-

cation of psychological testing emphasized the identification of intellectual deficiency and currently remains a prominent use of specific types of psychological tests. Educational problems provided a basis for additional areas of test development. The desire to classify students with reference to their ability to profit from different types of school instruction, the diagnosis of academic failures, counseling of high school and college students, the selection of aspirants for professional schools are a few of the educational uses of psychological testing (Anastasi, 1961).

The selection and placement of business and industrial personnel together with the certification of employees under Civil Service represent a recent and rapidly expanding use of psychological testing.

An examination of the Sixth Mental Measurement Yearbook (Buros, 1965) will attest to the rapid growth within the field of test development. On the basis of this proliferation, there is evidence supporting an additionally significant increase during the next decade.

The major exploitations of psychological testing have emphasized prediction and assessment. It is quite evident that the most pragmatic use of any selection device is to evaluate a

candidate's ability with a view toward adequate assessment and prediction. The historical and concurrent promise of psychological testing employes both of these rationales in fostering additional refinement within this area of psychological measurement. The refinement and improvement of techniques necessary for the future growth of psychological testing continues to be of paramount importance. No discipline, regardless of its philosophical basis, can survive under a pseudo-scientific attitude without a continual review of its guidelines and methods.

Fundamentally, a review of the literature pertaining to advanced testing techniques yields few new concepts concerning the assessment of the abilities of man. The preponderance of information presented in psychological literature deals with refined statistical methodology. While it is true that no scientist within the field would decry the necessary advances derived through statistical technology, it can be similarly argued that the concepts or testing devices which formulate the basis for the advanced statistics must be reviewed with a view toward improvement. While Gulliksen, (1950) in his treatise concerning the theory of mental tests, adequately presents the statistical foundation for psychological testing, he would be the first to admit that perhaps the concepts themselves should be

evaluated periodically. Other researchers in the field of psychometrics would probably reflect opinions of a similar nature.

The evolution of statistical techniques is quite rapid whereas the concurrent development of newer concepts to evaluate the complex processes of man is not nearly as pronounced. For example, little attention has been directed to assessing the cognitive factors of the mind. The current trend appears quite circular in its effect to create newer editions of the same examination based solely on refined statistics. Even though the complexity of assessment has been recognized for many years, the sharp ideological differences expressed by psychologists prohibit scientific investigation into areas such as "thinking." It is scientifically unpopular to consider an evaluation of man's capability in realms other than minutiae and practical skill applications. It is a sad commentary on a developing science to be satisfied with predictive efficiency coefficients of .35 based on an adequate sample size and to look no further in improving the technology involved in predicting this relationship other than to seek better statistical procedures.

It is interesting to note the development of psychological testing and review some of its historical foundations. In this regard, one notices quite readily that certain guiding principles within the field have little if any scientific foundation as

derived from formal experimentation, but have been developed on the basis of "accepted logic" rather than empirical investigation.

It is this author's firm belief that this growth of psychological testing necessitates a keener awareness of the associated problems, in addition to the full acceptance of the inherent responsibilities. The mere fact that testing has attained its current position of acceptance does not of itself indicate that such status is objectively warranted. Rather, it demonstrates the need for sound stratified research to substantiate empirically the basis of current test theory.

Since testing devices are utilized to gain some assessment of the individual participating in the examination, examiners would seemingly have the obligation of soliciting optimal performance from the examinee. In a teaching situation, utilizing a nonstandardized examination, the professor in preparing the test attempts to sample adequately the knowledge and skills that should be acquired on the basis of the imposed training or education program. The test constructor assumes that on the basis of the prepared examination, the students will be evaluated or assessed in proportion to their knowledge or acquired skill derived in the course. It is an implicit assumption that

learning can be demonstrated through the techniques of the examination. Holding all test construction premises constant, one would assume that if the examination was properly constructed, it should logically evaluate the area of knowledge it purports to measure.

If testing devices are ever to predict with a high degree of confidence or evaluate with minimal assurance, the optimal conditions underlying test performance must be isolated and eventually incorporated within the test situation. In essence, therefore, the purpose of this experimental study is an attempt to investigate one such condition which may influence test performance, that is, immediate knowledge of test achievement and to determine whether this has any consistent effect on objective examination results.

Of necessity, it must be the test constructor's goal to maximize the performance of all students subjecting themselves to the evaluation. This maximization of performance should more closely approximate what Gulliksen (1950) calls true test score.

$X_i = T_i + E_i$ or $E_i = X_i - T_i$

X_i = the score of the ith person on the test under consideration.

T_i = the true score of the ith person on this test.

E_i = the error component for the same person.

While it is stated that true examination score equals attained score plus some error, it is the obligation of the researchers to minimize the error contained in this formula. Recognizing the limitations of achieving a totally perfect evaluation where true score equals obtained score, it is nevertheless the responsibility of researchers involved in psychometrics to insure better guidelines to minimize existing error.

It is this author's belief that immediate knowledge of test performance may influence examination results and therefore should be investigated. The effort expended in this experimental project should, within the limits of the experimental design, present evidence concerning this factor.

CHAPTER II

REVIEW OF LITERATURE

The early investigators in the field of mental measurement- Galton (1883), Cattell (1888), Kraepelin (1895), Ebbinghaus (1897)- had focused their efforts on the study of individual differences from the standpoint of sensory and perceptual processes.

Binet and Henri (1895) criticized most of the available test series as being overly sensory in nature and therefore concentrating unduly on simple, specialized abilities. They argued that, in the measurement of the more complex functions, precision is not necessary, since individual differences are larger in these functions. They proposed a varied list of tests covering such functions as memory, imagination, attention, comprehension, suggestibility, aesthetic appreciation and many others (Anastasi, 1961).

Binet published "L'Etude Experimentale de L'Intelligence" (1903) in which he subjectively investigated his two daughters, Armande and Margueritte, on their ability to perform 20 given tasks. From investigations such as this into the analysis of

various tasks of intelligence, the Binet-Simon Scale emerged.

The Binet-Simon Scale (1905) contained 30 tests which were arranged in ascending order of difficulty. The difficulty level of the tasks was empirically determined by administering the tests to 50 normal children aged 3 to 11 years and also to some intellectually deficient children. The tests were designed to cover a wide variety of intellectual functions. They were tests of intelligence; but in 1905, Binet had only a vague idea of what he meant by intelligence (Varon, 1935). Binet scored the tests by adding age increments for each successfully accomplished task. In the Terman revisions, the number of tasks increased but the process remained essentially the same. The items indicated that the child who passed a test successfully possessed an ability that "....corresponds to the average ability of children of such and such an age" (Terman, 1937).

It was through this method of scoring, namely that of adding increments from unrelated scores and thereby obtaining a single total based on many diverse tasks, that Binet and subsequently Terman avoided the problem of absolute scaling. Currently, this problem is frequently overlooked, since it is customary to use Binet I.Q. results as total scores when assessing, for example, the placement of children within school programs. Terman was

aware of this problem and recommended the use of standard scores rather than I.Q.'s to indicate performance. Terman's faith in the adherence of his suggestions is apparent in his statement (Terman, 1937), "Whatever index of brightness is used, some will claim too much from it and others too little. The uninformed will read meaning into it which it does not connote and the over enthusiastic will, in too exclusive dependence upon it, ignore their lines of information which should be taken into account." As a result, Terman suggested that simple indexes be used and that training into the significance and limitations be adequately given.

From this brief overview of individual intelligence testing, and its costly procedures, it becomes somewhat apparent why the shift toward group testing emerged. It also is logical that any transition from individual to group testing would include as many of the acceptable procedural conditions of individual testing as possible. It is interesting, however, to note that on the basis of the extensive research conducted in the field of psychology dealing with knowledge of performance and results, few attempts, with the exception of programmed instruction, have been made to incorporate the established principles into the field of testing. For example, in reviewing related literature, it

became apparent that no well structured research project investigating the effects of immediate knowledge of test performance upon subsequent test results had been conducted.

Earlier investigations such as that conducted by Morgan and Morgan (1935) reviewed and investigated the problem of the effects of immediate knowledge of awareness of success and failure upon objective examination scores. The study attempted to evaluate the effects of immediate awareness of success and failure upon the results obtained from an objective examination. It was assumed in this experimental project that awareness of success or failure may produce no appreciable modification of results; it may cause increased effort, attention, critical observation and thereby improvement or it may prove discouraging and therefore detrimental.

In this study, an attempt was made to match the groups on their ability to perform on the Thurstone psychological examination. The matching results on this examination were somewhat less than accurate since the means and standard deviations of the two groups differed significantly. The experimental group, for example, tended to be superior in their performance on this examination. The authors, however, indicate that the difference between these groups was not "completely statistically significant."

The experimental condition consisted of using a self-scoring mimeographed copy of the examination. The self-scoring device afforded the candidates an opportunity to be aware of their overall success and failure in this examination. The retest intended to indicate true differences in favor of the experimental group. However, it should be pointed out that this difference also occurred during the matching examination. Therefore an immediate question arises as to whether the subsequent difference was a function of inadequate matching or truly significant on the basis of the experimental variable, i.e. knowledge of performance. Morgan's (1935) summary indicated that the self-scoring technique employed as a testing device appeared to produce no facilitating or detrimental effects. The authors further suggest, on the basis of the results found in their study, that the self-scoring technique suggested by Sidwell and Babcock (1933) appears to prove equally effective as a measuring device compared with the mimeographed form of the objective examination where success and failure is not apparent to the examinee. The authors conclude that immediate awareness of success and failure causes no significant change in the scores obtained from an objective examination.

It is this author's contention, however, that the results of Morgan's investigation should not be given too much credence. The weaknesses in both the experimental design and controls are apparent. Also the lack of statistical refinements in the analysis of the results suggest replication of this study before the results are accepted as fact.

Another early investigation of the effects of knowledge of results on learning and performance in a coordinated movement of two hands was conducted by J. L. Elwell and G. C. Grindley (1938). A number of similar experiments were conducted by various authors such as the experiments of Arps (1917), Crawley (1926) relative to the work done with arm and leg muscles, and Johanson (1922) dealing with reaction times and noting that knowledge of results can produce an improvement of overall performance. In many of Thorndike's experiments in the early 1930's on human learning (1931, 1932, and 1935) the subject had been placed in a situation in which he could make any one of a number of discrete responses and was then told whether his response was right or wrong. Thorndike studies in detail the way in which knowledge that a response is right leads to an increased tendency to perform that response in future actions. In his experiments, the term "right" can be equated with "reward" as used in early animal experi-

mentation on learning. Many of the foregoing investigations equated the overall effects of either reward or knowledge of results as an "incentive." Most authors, however, differentially viewed knowledge of results and reward, since it appeared that knowledge of results acted more as an incentive rather than an inhibitory factor.

Elwell and Grindley (1938) state that the quantitative results described in their paper suggested that knowledge of results in the acquisition of a human skill is similar to the effects of reward in animal learning. If, for example, a comparison of the results described in this paper were compared with maze learning experiments by animals, it would be noticed that there is no appreciable learning when no knowledge of results is given or when no reward is given but that learning occurs when knowledge of results or reward is given (Coleman, (1932; Grindley, 1932); and further that the acquired habit breaks up when knowledge of results is removed or when the reward is removed (Bruce, 1930; Grindley, 1932). The authors suggest that there is also an obvious parallel between the subject's attitude (i.e. keenness or desire to do well) in the experiment and "drive" in the animal study (Coleman, 1932).

Elwell and Grindley (1938) suggest that it would be interesting to attempt an explanation of the results of their experiments in terms of Thorndike's Law of Effect which has been used to explain animal learning. The authors further indicate, however, that in the acquisition of a muscular skill, such as that described in their paper, the learning cannot be regarded merely as the strengthening of the tendency to repeat movement which has been "rewarded" (by a high score). If a subject missed the bull's-eye he tried, on the next time, to correct this error by altering his response in the appropriate direction. In many of Thorndike's experiments (1931, 1932, and 1935), in which the subject is allowed to vary his behavior only between a limited number of discreet responses and is told whether he has made the "right response or wrong response," it may be legitimate to consider learning simply as the strengthening or weakening of tendencies to make each of these responses. But the authors stated that in experiments such as theirs, it is necessary to consider that knowledge of results, when the movement is not completely successful, introduces a tendency toward response replication. The authors call this the "directive effect" of knowledge of results.

In general, the authors cite several ways in which knowledge of results leads to improvement of performance in the experiments described in their paper. It should be further noted that in this investigation the removal of knowledge of results produced rapid deterioration in performance. The authors suggested these results seem to support the view that not only the acquisition of a skill, but also its maintenance depends upon continual "check-up" on the accuracy of the movement which has been made.

Using an apparatus in which the subjects attempted by a movement of two hands, to direct a spot of light onto the bull's-eye of a target, Elwell and Grindley found (a) that no improvement of accuracy of performance occurs without knowledge of results, (b) that improvement occurs with knowledge of results and (c) that removal of knowledge of results after the skill has been acquired leads to deterioration of performance.

Another related comprehensive study was conducted by the Psychological Laboratory at Cambridge under the authorship of MacPherson, Dees and Grindley entitled "The Effect of Knowledge of Results on Learning and Performance" (1948). This paper described an extension of previous work on the introduction and removal of visual knowledge of results to further motor skills most of which were intended to be "objectively simpler" from

those studied earlier. The tasks employed included such things as drawing a line of a certain length, exerting a given pressure on a lever or pressing a key for a given length of time. In each case the subject could be allowed to see the extent and direction of his error after every trial; but the apparatus was so arranged that this knowledge could be withheld. It was found that a continuous series of readings with visual knowledge of results produced more accurate performance.

In a follow-up to the 1948 research Valerie Dees and G. C. Grindley (1949) conducted a study on the effects of knowledge on learning and performance concerning the direction of error in very simple skills. The results of this and previous experiments by these authors show very clearly the importance of the direction as well as the amount of the error in any curve of learning or performance. Dees and Grindley hypothesized that when a subject is trying to repeat a movement (i.e. when he has hit the bull's-eye in the previous trial) he is trying to obtain proprioceptive sensations which match (Bartlett, 1932) his "memory trace" of those in the previous trial. Further, that the proprioceptive mechanism is susceptible to adaptation (Adrian, 1928), i.e. during the period immediately following a response the receptors or some more central part of the mechanism would be less sensitive.

The degree of adaptation would of course decline with the time since the last response. Further, that the "memory trace" left by the proprioceptive stimuli becomes less precise with the lapse of time, but there is no marked directional trend in such forgetting.

The hypothesis suggested by these authors makes no claim to explain the general mechanism of "learning" or the differences between what has been called the "incentive" and "directive" effects. They suggest that a complete theory should obviously link these findings with the many investigations of phenomena such as the "time error," muscular "after contraction" and positive and negative after images. But since these investigations do not yet form a coherent whole, the authors have not attempted to discuss them in this particular article but have confined themselves to the simple hypothesis about the present data. The general results presented by MacPherson, Dees and Grindley suggest that knowledge of results has its effects in improving objective performance on a variety of motor tasks, however, it can be clearly noted that the differential characteristics are not uniform and that both the learning and forgetting curves do not represent simple explanatory cycles.

Brown (1949) noted that although the belief in the efficiency of feedback is strongly entrenched in psychology there were actually few studies of the phenomenon as such. He noted that the majority of those in existence involved only two conditions (with and without feedback) and that there was almost no experimentation with systematic variations that have been introduced in the immediacy, continuousness, and specificity of feedback. Brown presented a proposed program of research to be conducted on psychological feedback in the performance of psychomotor tasks. He further suggests that feedback has three major functions any one of which may be maximized by appropriate manipulations of the learning situation (i.e. (1) feedback may provide specific information as to the extent and nature of errors which the subject can utilize in subsequent trials; (2) feedback may function principally as a reward or punishment for previously made responses with little or no specific information being given; (3) feedback may function to affect "the motivational level of the learner"). In general, Brown's work is noteworthy from the standpoint that it presents an early guideline for a proposed research program on psychological feedback (knowledge of results in the performance of psychomotor tasks).

Smode (1958) conducted a study on learning and performance in a tracking task under two levels of achievement information feedback. Smode's experiment was designed to provide an independent assessment of performance of effects and learning effects in a compensatory tracking task as a function of the method used in providing cumulative information as to achievement levels. A transfer of training design was employed, differentiating two basic conditions (high and low information feedback schedules) of the training phase into eight experimental subgroups in the transfer stage. On the transfer trials, one-half of the Ss continued to receive the same type of information imposed at the onset and one-half changed to the other schedule; one-half continued with the original target course while one-half transferred to the new target course equated for difficulty levels. The high and low information feedback conditions differed in terms of the aggregate effect of the following three parameters: (a) the amount of information presented; (b) the sensory mode of presentation; (c) the temporal characteristics of presentation. For both the time-on-target and the absolute integrated error scores, all groups showed consistent improvement over training trials; however, the high information feedback groups were superior at the end of the early trials and maintained this superiority throughout

the training period. Thus, the high level information feedback clearly facilitated performance. Comparisons of groups trained under different information feedback conditions and tested under identical conditions on transfer trials revealed significant differences which in all cases favored training under high level information feedback. The results of the experiments were identified as learning effects, although a carry-over of motivation hypothesis (i.e. persistence of a favorable attitude) was also considered a possible explanation by Smode. Smode further suggested that the subjective reactions to the tracking task indicated that interest level accrued as a function of increased information feedback and concluded that the effect of higher information feedback was mediated by an increase in motivation. He indicated that the manipulation of extrinsic information feedback may prove to be a useful technique for controlling human motivational levels in a variety of learning tasks.

Since 1915, S. L. Pressey (1950) has been researching problems related to immediate knowledge of performance. His inventions and research dealing with programmed instruction are well known. In an article published in 1950, Pressey summarized his work under four major conclusions:

1. He demonstrated a simple way to telescope into one

simultaneous process taking an examination, scoring it, informing students about their errors, and providing a method to find the right answers. By utilizing the special punch board examination sheet, the students were able to determine their performance in the examination. Pressey's purpose, however, in this experimentation was not to determine the effects of this immediate knowledge of performance on objective results but rather to develop a method that would be acceptable to examinees which would be both efficient and economical. Pressey notes that irrespective of any self-instructional values such a device may have, simple self-scoring devices should be worthwhile as a means of saving time and labor, and speeding up the total testing process.

2. The investigation showed that the new testing process, does transform test taking into a form of systematically directed self-instruction. Pressey notes that repetition of self-instructional tests brought marked reductions in the number of errors made.

3. His results further indicated that when the self-instructional tests were used systematically in college courses as an integral part of the teaching method, gains

were substantial and sufficiently generalized to improve understanding of the topic as a whole. Pressey's work suggested that punch board tests improved performance in regular classes as shown by higher scores on midterm and final examinations in comparison to comparable sections of the same course not employing the punch board. The device was found especially valuable with superior students.

4. It was noted that the punch board method was found simple and convenient for student use. This and other research conducted by Pressey suggested that human engineering can aid educational and training programs by test-teach devices of various types.

It appeared to this author, however, that Pressey's research as published in 1950 appeared more enumerative than statistical concerning the effectiveness of self-scoring examinations. While one would not question the worthwhile contribution presented by Pressey, it nevertheless appeared that his research suggested the need for empirical research to determine the varying effectiveness of this type of procedure on performance.

Angell (1949) conducted a study on the effects of knowledge of quiz results on final examination scores in freshman chemistry. His purpose was to determine the effects of immediate and delayed

knowledge of quiz results on three types of learning in freshman chemistry. Angell utilized Pressey's punch board method for students responses. The experimental group received knowledge of results by using the punch board during the midterm examination. Final examination scores were established as the criterion for this research project. Angell concluded that the differences between scores on the final examination were in favor of the experimental group that used the punch board and received immediate knowledge of results. The difference between the equated experimental and control group was significant at the one per cent level.

No review of the literature pertaining to knowledge of results would be complete without including the summary presented by Ammons (1956). In this publication, he presents a survey and tentative theoretical formulations concerning the area of effects of knowledge of performance. Ammons presents eleven generalizations based upon reasonably adequate studies concerning each of the factors. Some of the research cited by Ammons has already been included in the background material thus far presented in this research project. However, it is this author's belief that the generalizations formulated in this survey of literature are of significant value and must be included in any adequate

review of literature pertaining to the effects of knowledge of results. The generalizations presented by Ammons are as follows:

1. "The performer usually has hypotheses about what he is to do and how he is to do it, and these interact with knowledge of performance."

2. "For all practical purposes, there is always some knowledge of performance available to the human performer."

3. "A knowledge of performance affects the rate of learning and level reached by learning." Studies cited earlier by MacPherson, Dees and Grindley (1948), Pressey (1950), Angell (1949) and Morgan (1935) support this generalization.

4. "Knowledge of performance affects motivation." Helmstadter and Ellis (1952) tried various kinds of goal setting procedures with a block turning task and concluded from the results that simple knowledge of performance led to increased motivation. Pressey (1950) found that students who had used a punch board device for self-scoring of quiz results, preferred this way of taking multiple choice tests. Book and Norvelle (1922) noted experimentally that motivation resulted from knowledge of performance and that the subjects included in the experiment

eventually developed ways of keeping track of their scores and also accrued greater interest in the tasks to be completed. As mentioned earlier, MacPherson, Dees and Grindley (1938) noticed that giving knowledge of performance led to a more favorable general attitude toward the experiment.

5. "The more specific the knowledge of performance, the more rapid the improvement and the higher the level of performance." Trowbridge and Cason (1932) investigated the problem of specificity of knowledge of performance and its effects on individuals. Waters (1933) found that improvement was "roughly proportional to the degree of information given" about the correctness of the estimates in the experiment. Bilodeau and Morin (1951) demonstrated with a "pedestal sight manipulation test" and noted that the trainees made better scores with more specific information concerning their original performance.

6. "The longer the delay in giving knowledge of performance, the less effect the given information has." A number of studies support this generalization. Saltzman (1951) reported slower maze learning when knowledge of performance was withheld for six seconds. Keller's (1943) work in code reception indicated that performance knowledge after

each test word was more effective than knowledge of performance given only after much larger units.

7. "In the case of discontinuous tasks where knowledge of performance is given, small intervals between trials are generally better for learning than are larger ones."

8. "When knowledge of performance is decreased, performance drops."

9. "When knowledge of performance is decreased, performance drops more rapidly when trials are relatively massed."

10. "When subjects are not given supplementary knowledge of performance by the experimenter, the subjects that maintain a performance level probably have developed some substitute knowledge of performance."

11. "When direct (supplementary) knowledge of performance is removed, systematic 'undershooting' or 'overshooting' may appear in motor tasks." Eaton's work (1935) and the study reported by Dees and Grindley (1951) support this generalization.

While Ammon's work is certainly comprehensive from the standpoint of surveying literature and establishing guidelines about knowledge of performance, it must be noted that several of the generalizations do not have sufficient supporting empirical

information to justify unequivocal acceptance.

Rethlingshafer (1963) suggests in her text on motivation that learning may be simple or complex bringing only a slight modification or complete change and that learning as broadly defined is information. Experiments studying the effects of knowledge of results cited in the literature vary the degree of information ranging from incomplete to exact knowledge needed. Rethlingshafer notes that it is important to remember that achievement information feedback tells the subject how the results conformed to some norm, including possibly the subject's own standard of performance. Knowledge of achievement may have an inciting effect on behavior. However, it is not clear whether the more rapid improvement resulting from feedback during the early practice trials on a task is indicative of more rapid learning, or only of a heightened level of performance attributable to greater motivation and effort. The complexity of this problem cannot, however, be answered on the basis of single dimensional analysis.

It seems clear from the review of literature that there is precedent for the investigation suggested in this experimental project. It is similarly true that no definitive results of statistical significance have been presented in the literature

to confirm or deny the hypothesis that knowledge of performance in a testing situation facilitates uniform improvement by the examinee which is the defined problem of this research paper.

CHAPTER III

METHOD

The present study was undertaken to investigate the effects of immediate knowledge of performance on test results. It was intended to determine whether there was any significant difference between performance of students receiving immediate knowledge of results in contrast to an equated group of subjects that did not receive such knowledge of performance.

In order to fulfill the requirements of an adequate experimental design, the author found it necessary to use a standardized, general ability examination on which parallel forms and research data were available. As an integral part of this study, one standardized examination included in the Dental Aptitude Testing Program was selected. The School and College Abilities Test (SCAT), Forms 1B and 1D, was chosen as the testing device to be incorporated within the experimental design of this study. This examination was selected as the principal testing device, in order to gain an assessment relative to the problem of immediate knowledge of performance and whether any significant differing

effects occur on verbal versus quantitative aspects of general ability.

The School and College Abilities Test (SCAT) was devised by Educational Testing Service in 1955 primarily as an aid in estimating the capacity of students in grades 4-14 to undertake additional education. The general abilities measures as derived through the use of the SCAT examination have been included in the Dental Aptitude Testing Program since October, 1959 as a replacement for the ACE psychological examination.

The School and College Abilities Test is designed for group administration and easily scored on either an IBM 805 or 1230 scoring machine. Two special modified versions of these examinations (Forms 1B and 1D) are utilized for this experimental project with a time limit of 60 minutes.

Since the main condition of this experimental research project is to determine the effects of immediate knowledge of test performance on simultaneous or concurrent test results, it was necessary for this author to construct a special answer sheet. The special answer sheet permitted the experimental group to receive immediate knowledge of item performance whereas the control group used a similar answer sheet in which no feedback of results was obtained. The specially prepared answer sheet

was adapted from the IBM 805 answer sheet format. A sample of the answer sheets developed for this experimental project appears in the Appendix.

The 388 students included in this experimental study were recently enrolled first year dental students at the following schools:

			CONTROL GROUP	EXPERIMENTAL GROUP
Loyola University	N =	86 =	43	43
Northwestern University	N =	68 =	34	34
University of Illinois	N =	88 =	44	44
University of Michigan	N =	96 =	48	48
Washington University	N =	50 =	25	25
	T =	388 =	194	194

The sample of students was chosen for this experimental study because of the availability of prior test scores and other background information pertaining to the general characteristics of each of these first year dental students.

All of the subjects participated in the pretest condition as part of their admission to dental school. Therefore, prior Dental Aptitude Test and SCAT scores are available for all of the students included in the sample. Approximately one year

intervened between the pre and post test condition.

The 388 subjects were administered one form of the SCAT examination under the routine administration pattern. The subjects were then randomly assigned into an experimental (N = 194) and control group (N = 194) by randomly distributing the two special answer sheets. The groups were not specifically matched. The experimental condition utilized the principle of immediate awareness of item test performance on an alternate form of the SCAT examination. In this instance, the independent or treatment variable was the varied condition of receiving immediate feedback of performance when the experimental group attempted each item. This feedback took the form of printed letters (i.e. C or I signifying correct or incorrect) on the answer sheet placed below a silk screen, therefore, not visible to the candidate until he attempted the item by erasing the silk screen. The dependent or criterion variable was the number of correct items on the examination after the imposed experimental condition.

The learning carry-over from the pretest was controlled by the administration of an alternate, parallel test form of the SCAT examination. The specially prepared answer sheet permitted the immediate feedback of information and was the ~~major~~ experimental condition imposed.

LEWIS TOWERS
LOYOLA
UNIVERSITY

Differences in performance between the experimental and control group as affected by the experimental condition were analyzed. Both within and between group analyses were performed on the basis of the overall design of this experiment.

CHAPTER IV

RESULTS

The sample of students was selected from the September, 1966 first year classes at five dental schools. Each first year dental class was divided in half and the subjects were randomly assigned to the experimental and control group. There were 388 students utilized as the sample for this experiment. (Control Group - N = 194, Experimental Group - N = 194.)

In addition to the overall availability of information concerning these two groups, certain additional data were available concerning the comparability of the experimental and control group. For example, Dental Aptitude Test scores were available for each of the 388 students. As can be noted from Table I, the mean academic average for the control group was 4.87 on a coded score basis (range -1 to 9). Similarly, it can be noted that the mean academic average for the experimental group was 4.93. The academic average score as obtained by the Dental Aptitude Testing Program is a composite measure of scores derived from the general abilities test (SCAT), the basic science examination in

biology and chemistry and a reading comprehension test. The academic average is a simple composite of these examinations. It is apparent that no significant difference exists between the control and experimental group on the ability scores as measured by this portion of the Dental Aptitude Test. It should be noted, however, that these students were not matched but drawn from a homogeneous population and randomly assigned to the control and experimental group.

Since the major experimental condition utilized the SCAT examination as the dependent variable, it was most essential that the control and experimental group did not differ on their pretest performance on this examination. It can be noted from a review of Table I that the mean performances of the control and experimental group on the SCAT pretest examination yielded no significant difference. The quantitative, linguistic and total test scores did not deviate by more than 1/100th of a percentage point. It is apparent from the review of these data that the random assignment of subjects and the large sample produced a well-balanced and matched experimental and control group for this study.

Another factor considered in the matched comparison of the experimental and control group was the number of years of pre-professional education completed at the time of this experiment.

As can be noted from Table I, the mean number of years of predental education for the control group was 2.84 whereas it was 2.91 for the experimental group. Again, no significant difference existed between the number of years of preprofessional education.

TABLE I

COMPARISON OF EXPERIMENTAL AND CONTROL GROUP ON DENTAL APTITUDE TEST PERFORMANCE, SCAT PRETEST PERFORMANCE, AGE AND NUMBER OF YEARS OF PREDENTAL EDUCATION

	Control Group N = 194	Experimental Group N = 194
	Dental Aptitude Test Performance	
Academic Average		
Mean	4.87	4.93
S.D.	1.07	1.12
Manual Average		
Mean	5.01	5.03
S.D.	1.28	1.26
	SCAT Pretest Performance	
Quantitative		
Mean	4.86	4.86
S.D.	1.80	1.99
Linguistic		
Mean	4.71	4.70
S.D.	1.68	1.65
Total		
Mean	4.88	4.87
S.D.	1.58	1.67
	Number of Years of Predental Education	
Mean	2.84	2.91
	Age	
Mean	22.14	22.22

TABLE II

CONTROL GROUP
TOTAL SCAT MEAN SCORES
BY DENTAL SCHOOLS

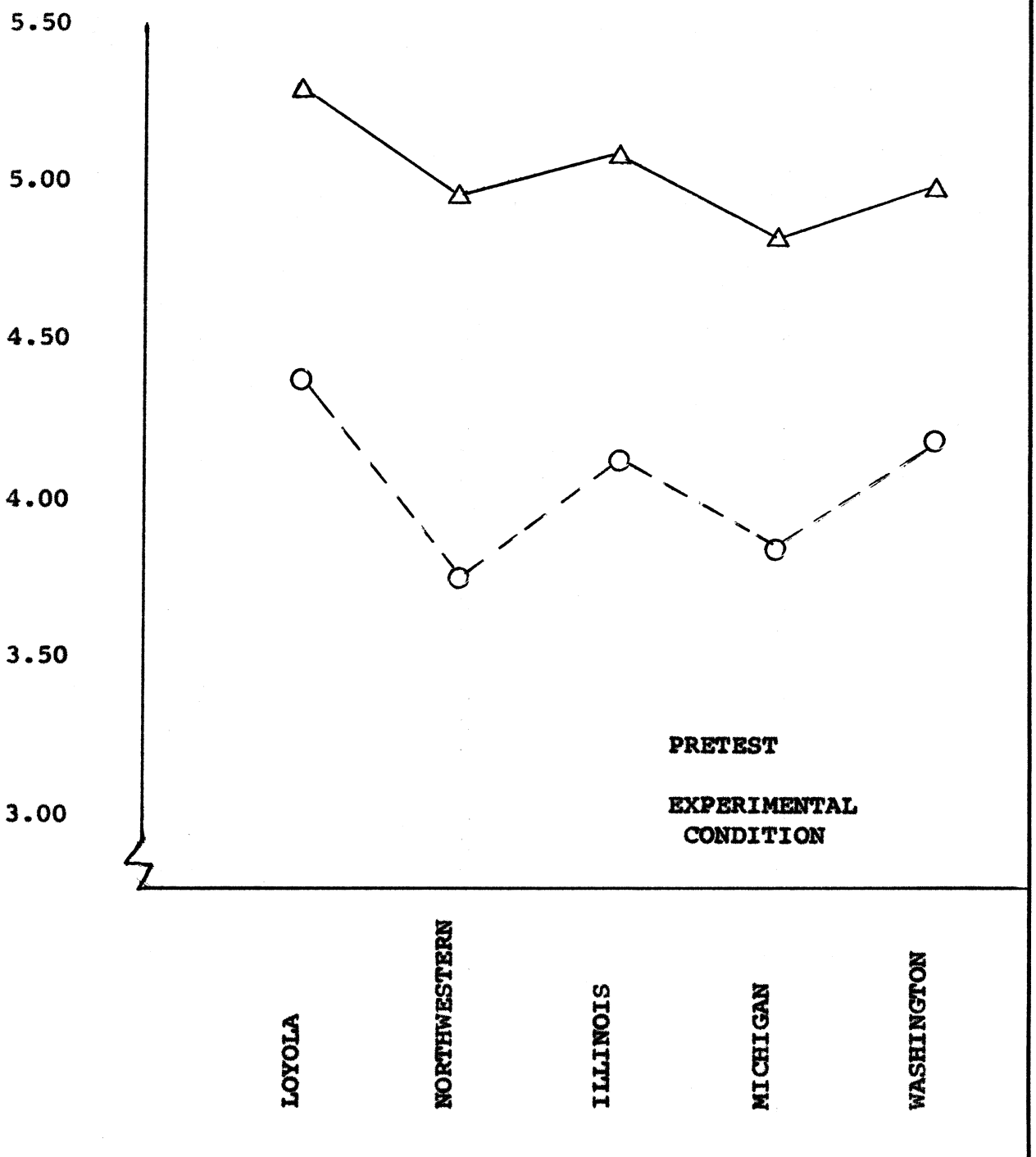


TABLE III

EXPERIMENTAL GROUP
TOTAL SCAT MEAN SCORES
BY DENTAL SCHOOLS

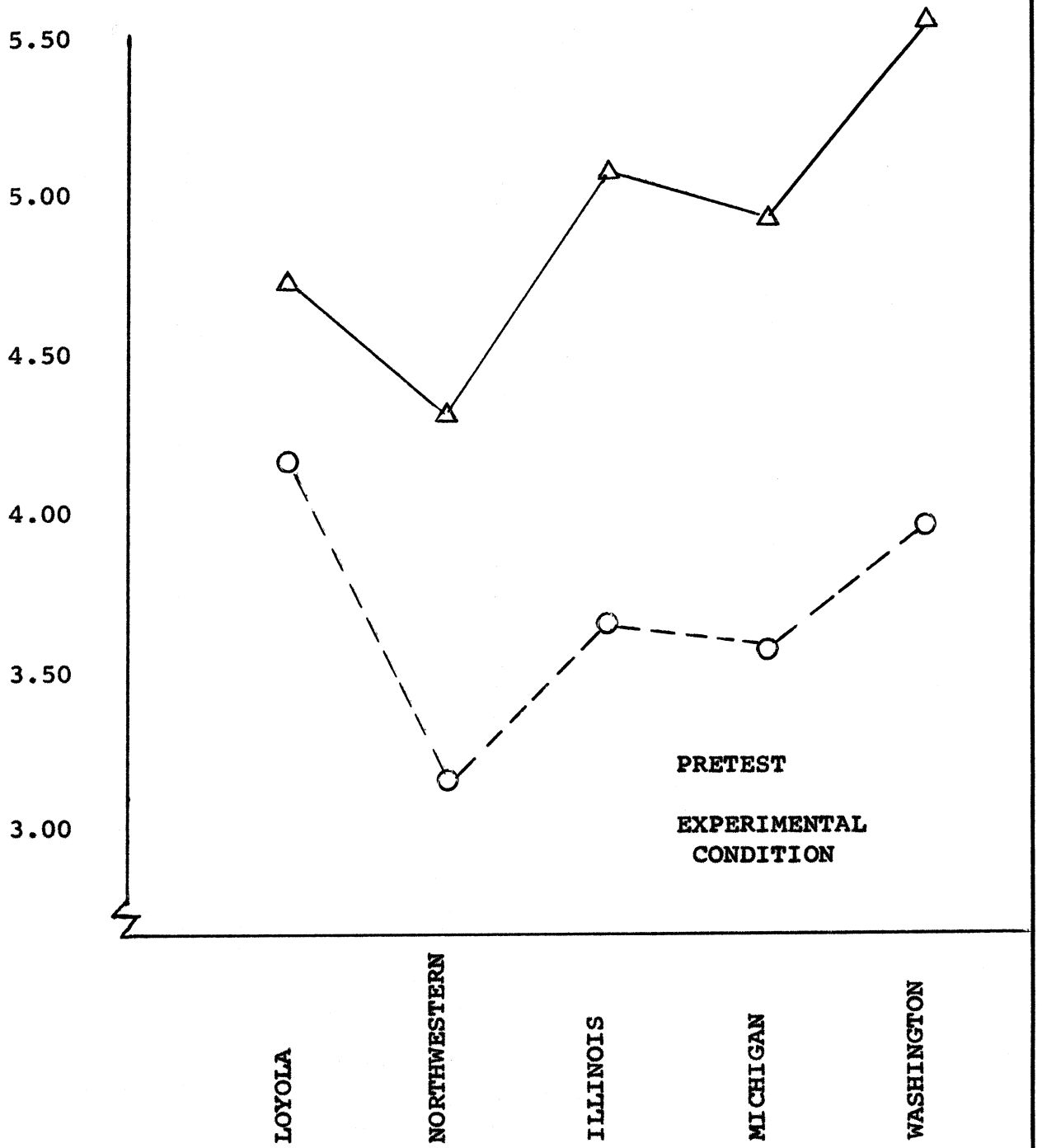


TABLE IV

COMPARISON OF PRE AND POST TEST CONTROL GROUP PERFORMANCE ON SCAT QUANTITATIVE TEST ITEMS

	Pretest N = 194	Post Test N = 194
Mean	4.86	3.34
S.D.	1.80	1.95
Variance	3.24	3.80
σ_M	.130	.140

$r = .58$

$\sigma_{d_M} = .1240$

$D_M = 1.52$

$\overline{z} = 12.26^*$

* = $P < .01$

TABLE V

COMPARISON OF PRE AND POST TEST CONTROL GROUP PERFORMANCE ON SCAT LINGUISTIC TEST ITEMS

	Pretest N = 194	Post Test N = 194
Mean	4.71	4.61
S.D.	1.68	1.77
Variance	2.82	3.13
σ M	.121	.127

$r = .68$

$\sigma_{d_M} = .0993$

$D_M = .10$

$\overline{Z} = 1.007$

TABLE VI

COMPARISON OF PRE AND POST TEST CONTROL GROUP PERFORMANCE ON SCAT TOTAL TEST ITEMS

	Pretest N = 194	Post Test N = 194
Mean	4.88	4.05
S.D.	1.58	1.70
Variance	2.50	2.89
σ_M	.114	.122

$r = .69$

$\sigma_{d_M} = .0932$

$D_M = .83$

$\overline{z} = 8.91^*$

* = $P < .01$

TABLE VII

COMPARISON OF PRE AND POST TEST EXPERIMENTAL GROUP PERFORMANCE ON SCAT QUANTITATIVE TEST ITEMS

	Pretest N = 194	Post Test N = 194
Mean	4.86	2.83
S.D.	1.99	2.03
Variance	3.96	4.12
σ_M	.143	.146

$r = .57$

$\sigma_{d_M} = .1340$

$D_M = 2.03$

$\overline{z} = 15.15^*$

* = $P < .01$

TABLE VIII

COMPARISON OF PRE AND POST TEST EXPERIMENTAL GROUP PERFORMANCE ON SCAT LINGUISTIC TEST ITEMS

	Pretest N = 194	Post Test N = 194
Mean	4.70	4.49
S.D.	1.65	1.86
Variance	2.72	3.46
σ_M	.119	.134

$r = .70$

$\sigma_{d_M} = .099$

$D_M = .21$

$\overline{Z} = 2.12^{**}$

** = $P < .05$

TABLE IX

COMPARISON OF PRE AND POST TEST EXPERIMENTAL GROUP PERFORMANCE ON SCAT TOTAL TEST ITEMS

	Pretest N = 194	Post Test N = 194
Mean	4.87	3.66
S.D.	1.67	1.91
Variance	2.79	3.65
σ_M	.120	.138

$r = .69$

$\sigma_{d_M} = .237$

$D_M = 1.21$

$\overline{Z} = 5.11^*$

* = $P < .01$

TABLE X

COMPARISON OF CONTROL AND EXPERIMENTAL GROUP PERFORMANCE ON SCAT QUANTITATIVE TEST RESULTS DURING THE POST TEST CONDITION

	Control Group N = 194	Experimental Group N = 194
Mean	3.34	2.83
S.D.	1.95	2.03
Variance	3.80	4.12
σ_M	.140	.146

$\sigma_{d_M} = .202$

$D_M = .510$

$\overline{Z} = 2.52^{**}$

** = $P < .05$

TABLE XI

COMPARISON OF CONTROL AND EXPERIMENTAL GROUP PERFORMANCE ON SCAT LINGUISTIC TEST RESULTS DURING THE POST TEST CONDITION

	Control Group N = 194	Experimental Group N = 194
Mean	4.61	4.49
S.D.	1.77	1.86
Variance	3.13	3.46
σ_M	.127	.134

$$\sigma_{d_M} = .184$$

$$D_M = .120$$

$$\overline{Z} = .652$$

TABLE XII

COMPARISON OF CONTROL AND EXPERIMENTAL GROUP PERFORMANCE ON SCAT TOTAL TEST RESULTS DURING THE POST TEST CONDITION

	Control Group N = 194	Experimental Group N = 194
Mean	4.05	3.66
S.D.	1.70	1.91
Variance	2.89	3.65
σ_M	.12	.14

$\sigma_{d_M} = .184$

$D_M = .39$

$\overline{Z} = 2.12^{**}$

** = $P < .05$

TABLE XIII

COMPARISON OF CONTROL AND EXPERIMENTAL GROUP ON THE UNATTEMPTED ITEMS IN THE EXPERIMENTAL CONDITION

THE KOLMOGOROV-SMIRNOV TWO-SAMPLE TEST

	Control Group N = 194	Experimental Group N = 194
Mean	3.80	5.51

$$D = \text{maximum}/Sn_1(x) - Sn_2(x)/$$

$$D = .154^{*}$$

$$.025 \text{ level} = 1.48 \sqrt{\frac{n_1 + n_2}{n_1 \, n_2}} = 1.48 \sqrt{\frac{388}{37638}}$$

$$= .148$$

$$.01 \text{ level} = 1.63 \sqrt{\frac{n_1 + n_2}{n_1 \, n_2}} = 1.63 \sqrt{\frac{388}{37638}}$$

$$= .163$$

* = $\therefore$ .154 is significant beyond the .025 level

TABLE XIV

COMPARISON OF CONTROL AND EXPERIMENTAL GROUP PERFORMANCE ON PART I (LINGUISTIC ITEMS # 1-30) RAW SCORE TEST RESULTS DURING THE POST TEST CONDITION

	Control Group N = 194	Experimental Group N = 194
Mean	23.00	22.83
S.D.	3.67	3.94
Variance	13.47	15.52
σ_M	.264	.283

$$\sigma_{d_M} = .387$$

$$D_M = .170$$

$$\overline{Z} = .439$$

TABLE XV

COMPARISON OF CONTROL AND EXPERIMENTAL GROUP PERFORMANCE ON PART II (QUANTITATIVE ITEMS # 31-55) RAW SCORE TEST RESULTS DURING THE POST TEST CONDITION

	Control Group N = 194	Experimental Group N = 194
Mean	21.11	21.07
S.D.	3.10	2.96
Variance	9.61	8.76
σ_M	.223	.213

$\sigma_{d_M} = .308$

$D_M = .040$

$\overline{Z} = .130$

TABLE XVI

COMPARISON OF CONTROL AND EXPERIMENTAL GROUP PERFORMANCE ON PART III (LINGUISTIC ITEMS # 56-85) RAW SCORE TEST RESULTS DURING THE POST TEST CONDITION

	Control Group N = 194	Experimental Group N = 194
Mean	23.36	22.86
S.D.	3.71	3.82
Variance	13.76	14.59
σ_M	.267	.275

$$\sigma_{d_M} = .383$$

$$D_M = .50$$

$$\overline{Z} = 1.305$$

TABLE XVII

COMPARISON OF CONTROL AND EXPERIMENTAL GROUP PERFORMANCE ON PART IV (QUANTITATIVE ITEMS # 86-110) RAW SCORE TEST RESULTS DURING THE POST TEST CONDITION

	Control Group N = 194	Experimental Group N = 194
Mean	17.36	15.74
S.D.	5.08	6.22
Variance	25.81	38.69
σ_M	.365	.447

σ_{d_M} = .577

D_M = 1.620

$\overline{Z}$ = 2.808*

* = $P < .01$

TABLE XVIII

COMPARISON OF CONTROL AND EXPERIMENTAL GROUP AMONG LOW POST TEST SCAT PERFORMANCES

	Control Group N = 35	Experimental Group N = 55
	Low = -1 to 2	
Pretest		
Mean	3.20	3.58
S.D.	1.14	1.64
Variance	1.300	2.690
M	.196	.223
	σd_M = .297	
	D_M = .380	
	$\overline{Z}$ = 1.279	
Post Test		
Mean	1.40	1.31
S.D.	.76	.73
Variance	.578	.533
M	.130	.099
	σd_M = .163	
	D_M = .090	
	$\overline{Z}$ = .552	
r	.10	.14

TABLE XIX

COMPARISON OF CONTROL AND EXPERIMENTAL GROUP AMONG MIDDLE POST TEST SCAT PERFORMANCES

	Control Group N = 127	Experimental Group N = 108
	Middle = 3 to 5	
Pretest		
Mean	4.88	4.99
S.D.	1.29	1.33
Variance	1.664	1.769
M	.115	.129
	σd_M = .173	
	D_M = .110	
	$\overline{Z}$ = .636	
Post Test		
Mean	4.14	4.03
S.D.	.80	.76
Variance	.640	.578
M	.071	.074
	σd_M = .103	
	D_M = .110	
	$\overline{Z}$ = 1.068	
r	.38	.45

TABLE XX

COMPARISON OF CONTROL AND EXPERIMENTAL GROUP AMONG HIGH POST TEST SCAT PERFORMANCES

	Control Group N = 32	Experimental Group N = 31
	High = 6 to 9	
Pretest		
Mean	6.69	6.74
S.D.	1.00	1.47
Variance	1.000	2.161
M	.180	.268
	σd_M = .323	
	D_M = .050	
	$\overline{Z}$ = .155	
Post Test		
Mean	6.56	6.58
S.D.	.85	.80
Variance	.723	.640
M	.153	.146
	σd_M = .211	
	D_M = .020	
	$\overline{Z}$ = .095	
r	.58	.47

TABLE XXI

EXPERIMENTAL CONDITION COMPARISON OF CONTROL AND EXPERIMENTAL GROUP PERFORMANCES AMONG LOW RANGE PRETEST SCAT PERFORMANCES

	Control Group N = 15	Experimental Group N = 14
	Low = -1 to 2	
Pretest		
Mean	1.93	1.79
S.D.	.28	.39
Post Test		
Mean	1.87	1.50
S.D.	1.19	1.40
Variance	1.416	1.960
σ M	.318	.388
	$\sigma\, d_M$ = .502	
	D_M = .370	
	$\overline{z}$ = .737	
r	.36	-.09

TABLE XXII

EXPERIMENTAL CONDITION COMPARISON OF CONTROL AND EXPERIMENTAL GROUP PERFORMANCES AMONG MIDDLE RANGE PRETEST SCAT PERFORMANCES

	Control Group N = 114	Experimental Group N = 115
	Middle = 3 to 5	
Pretest		
Mean	4.26	4.21
S.D.	.80	.77
Post Test		
Mean	3.55	3.09
S.D.	1.37	1.55
Variance	1.877	2.403
σ M	.129	.145
	σd_M = .194	
	D_M = .460	
	$\bar{z}$ = 2.371**	
r	.38	.35

** = P < .05

TABLE XXIII

EXPERIMENTAL CONDITION COMPARISON OF CONTROL AND EXPERIMENTAL GROUP PERFORMANCES AMONG HIGH RANGE PRETEST SCAT PERFORMANCES

	Control Group N = 65	Experimental Group N = 65
	High = 6 to 9	
Pretest		
Mean	6.63	6.71
S.D.	.85	.99
Post Test		
Mean	5.42	5.15
S.D.	1.32	1.50
Variance	1.742	2.250
σ M	.165	.188
	σd_M = .250	
	D_M = .270	
	$\overline{Z}$ = 1.080	
r	.44	.54

CHAPTER V

DISCUSSION

Tables II and III present a graphic illustration by dental school of the total SCAT mean scores in both the pretest and experimental test condition. As can be noted, the general performance on the post test condition for both the experimental and control group was significantly less than the performance obtained during the pretest condition. Each of the dental school classes exhibited a similar reduction in overall performance scores during the second administration of the examination. This reduction occurred in both the control and experimental group performances.

A within group analysis was conducted on the experimental and control groups incorporated in this study. The purpose of this analysis was to determine the significant difference in performance on the pre and post test SCAT examination for the quantitative, linguistic and total test scores. Tables IV, V, VI, VII, VIII and IX present the detailed analysis of these findings. Tables IV, V and VI present the within group analysis of the control group whereas Tables VII, VIII and IX present a similar

analysis for the experimental group. The mean for the original or pretest condition, in addition to the post test or experimental examination results, was compared for significant differences. With the exception of the linguistic section of the SCAT examination, all of the within group analyses demonstrated significant differences in performance in both the experimental and control groups. This indicated that the second examination or experimental condition resulted in significantly lower mean performances for the 194 subjects contained in each group. The rationale for this reduction is twofold and is discussed later in the analysis of the results.

As mentioned in the procedure, each of the subjects included in this research project were examined on alternate parallel forms of the SCAT examination. However, it can be noted by observing the correlations appearing in Tables IV-IX that each of these correlations represent values considerably less than would normally be obtained in a test-retest reliability situation. Educational Testing Service, in its technical manual concerning the SCAT examination, indicates that the reliability between forms approximates .90 to .95. Why is the correlation for this size sample supressed? The explanation for this lower than expected test-retest correlation is based on the restriction of range

phenomenon. The reader must remember that the SCAT examination was part of the preselection or screening used for admission to dental school. The subjects included in this study were selected on the basis of this criterion examination. Therefore, the total range of available scores for the first year class on this examination was reduced by approximately 25 per cent. As mentioned earlier, the range of coded scores is a -1 to a +9. Students typically accepted for admission to dental school on the basis of this or the other variables included in the aptitude examination usually achieve a 4 coded score as a minimum. Therefore, the greatest proportion of the 388 subjects included in this sample was selected above that norm. The correlations as indicated in Tables IV-IX represent zero-order correlations and have not been adjusted for restriction in range. Sufficient evidence exists concerning the comparability between SCAT Forms 1B and 1D and therefore the condition of parallel examinations did exist.

The between group analysis (i.e. experimental and control group) was conducted and appears in Tables X, XI and XII. The purpose was to determine the effects of immediate knowledge of performance on concurrent test results. It should be noted that the control group did not receive immediate knowledge of item performance whereas through the use of the special answer sheet

the experimental group did receive immediate feedback. On the basis of related literature, one might hypothesize that the experimental group receiving immediate knowledge of performance should improve or be superior to the control group in overall results. A review of Tables X, XI and XII clearly demonstrates that such was not the case. While a significant performance difference existed on two sections of the SCAT examination (i.e. quantitative and total test scores) in favor of the control group, a further analysis of the results was indicated.

While the SCAT examination would not be considered a speed test as such, a moderate time limit of 60 minutes was imposed. In the routine administration of this examination during the pre-test condition, this time limit was sufficient since almost all candidates completed the test. However, by using a different type of answer sheet in this experimental project, the time limit of 60 minutes appeared too constraining for the majority of the candidates as can be noted from the mean performances obtained in the second administration. It is obvious that the mean performances of all students were significantly lower during the second administration.

The SCAT examination is arranged in such a way that the quantitative and linguistic items are split in quartiles.

	ITEM NUMBERS	
Part I	1 - 30	Linguistic
Part II	31 - 55	Quantitative
Part III	56 - 85	Linguistic
Part IV	86 - 110	Quantitative

An analysis of the unattempted items was of paramount importance to determine whether there was any significant difference in this aspect between the control and experimental group. It should be noted that the last 25 items included in this examination represent one-half of the quantitative section of the test.

Table XIII compared the control and experimental group on the number of unattempted items in the experimental condition by means of the Kolmogorov-Smirnov Two Sample Test. Table XIII clearly notes that the mean average number of unattempted items for the experimental was 5.51 whereas the mean unattempted items for the control group was only 3.80. The results of the two-sample test indicated that the chance probability of this occurrence was significant beyond the .025 level. On the basis of this important finding, a reinterpretation of Tables X, XI and XII was indicated. Since there is no significant difference between the experimental and control group on the linguistic or verbal section of the SCAT examination and since sub-

jects completed all of the items contained in this section, it is reasonable to assume that on the basis of the results presented in Table XIII the significant difference appearing on the quantitative section of the SCAT examination is in part due to the inability of the experimental group to complete the last number of items. The control group completed the examination more often than the experimental group. In addition, since the total SCAT score is dependent upon both the quantitative and linguistic items, the fact that a large percentage of the quantitative items was not completed by the experimental group would account in large measure for the significant difference in performance between the control and experimental group on total SCAT scores.

As mentioned earlier, the SCAT examination was divided into four sections. An analysis of each section was prepared to determine if true differences existed on the other parts of the test. Tables XIV through XVII present the results obtained by performing a test section analysis comparing the control and experimental group. The results indicated that no significant difference existed between control and experimental group performance on

Part I	Item 1 - 30	Linguistic
Part II	Item 31 - 55	Quantitative
Part III	Item 56 - 85	Linguistic

Part IV (Quantitative Items 86 - 110) did produce significant differences in performance which are reflected by the results in Table XVII. This was consistent with the other findings in this study and resulted from the number of unattempted items included in this last section.

Tables XVIII through XXIII present comparisons of low, middle and high performers on both the pretest and experimental condition. A between and within group analysis was completed to note any significant differences between these categories. The analysis of the data both on the basis of within and between group statistics yielded no significant differences in the results other than those noted on the prior tables.

In addition, the author prepared a detailed analysis on a school-by-school basis. This analysis included a test section comparison of the four subparts of the examination. The results of this comparison supported the overall results of this experimental project. All of the school results were consistent with the findings noted in the total group analysis. The results failed to disclose any significant difference between the five schools included in this study.

On the basis of these findings, it is the conclusion of this author that knowledge of results in a testing situation using a

standardized, general abilities examination constructed of discreet items is neither a positive nor detrimental factor on overall test performance. It should be noted, however, that this conclusion is based on the part analysis of the examination results and not on the basis of total test scores. In general, it would appear that knowledge of performance in a routine standardized testing situation had no consistent effect on objective examination test performance.

CHAPTER VI

SUMMARY

The purpose of this research project was to investigate the effects of immediate knowledge of test performance on concurrent test results. It was intended to determine whether there was any significant difference between performance of students receiving immediate knowledge of item results in contrast to an equated group of subjects that did not receive such knowledge of performance.

Three hundred and eighty-eight first year dental students enrolled in five universities were randomly divided into an experimental and control group for this study. As an integral part of this study, the School and College Abilities Test was chosen as the testing device for this experimental study. In addition, a special self-scoring answer sheet was prepared by the author to permit the availability of immediate knowledge of item performance to the experimental group. A similar answer sheet was used with the control group without the presentation of immediate knowledge of performance. All of the subjects participated in

the pretest condition as part of their admission to dental school, therefore prior Dental Aptitude Test and SCAT scores were available for the sample.

In this instance, immediate knowledge of performance took the form of printed letters (i.e. C or I signifying correct or incorrect) on the answer sheet placed below a silk screen and not visible to the candidate until he attempted the item by erasing the silk screen. The dependent variable was the number of correct items on the examination after the imposed experimental condition.

An analysis of the results focused on significant differences between the two groups on the basis of performance in the experimental or control condition of the examination. Significant differences in performance were obtained between the control and experimental group on both the quantitative and total test scores included in the SCAT examination. However, no significant difference in performance appeared on the linguistic or verbal section of the examination. A further detailed four part test section analysis revealed that this significant difference was attributed to the abbreviated time period which did not permit the candidate ample time to complete the examination under the experimental condition. Therefore, their overall mean performances during the experimental condition were lower. The experimental

condition permitting knowledge of item performance required that applicants take a longer period of time to answer the questions.

It was concluded, however, that the overall performance of subjects receiving immediate knowledge of results in a testing situation did not produce facilitating nor detrimental effects on the outcome of concurrent examination results. While the related literature would tend to suggest improved performance occurring on the basis of analogous situations in other fields, such results did not occur in this investigation. Since the SCAT examination was comprised of mutually exclusive or discreet items having little, if any, relationship to the prior items, test performance was not enhanced by receiving immediate knowledge of results for each item.

BIBLIOGRAPHY

Ammons, R. B. Effects of knowledge of performance: A survey and tentative theoretical formulation. J. Gen. Psychol., 1956, 54, 279-299.

Ammons, R. B. Effect of distribution of practice on rotary pursuit "hits." J. Exper. Psychol., 1951, 41, 17-22.

Ammons, R. B. An analysis of "hits" in continuous rotary pursuit before and after a single rest. J. Gen. Psychol., 1953, 48, 3-10.

Ammons, R. B. Experiential factors in visual form perception: I. Review and formulation of problems. J. Gen. Psychol., 1954, 84, 3-25.

Ammons, R. B. Knowledge of performance: Survey of Literature, some possible applications, and suggested experimentation. Wright-Patterson Air Force Base, Ohio: Wright Air Development Center, Aero Medical Laboratory, February, 1954. (WADC Technical Report 54-14.)

Anastasi, Anne. Psychological Testing. New York: The Macmillan Co., 1961.

Angell, G. W. The effect of immediate knowledge of quiz results on final examination scores in freshman chemistry. J. Educ. Res., 1949, 42, 391-394.

Arps, G. F. A preliminary report on work with knowledge versus work without knowledge of results. Psychol. Rev., 1917, 24, 449-455.

Arps, G. F. Work with knowledge of results versus work without knowledge of results. Psychol. Monog., 1920, 28, No. 125.

Atkinson, R. C. & Ammons, R. B. Experiential factors in visual form perception: II. Latency as a function of repetition. J. Exper. Psychol., 1952, 43, 173-178.

Bilodeau, E. A., & Morin, R. E. A preliminary investigation of the effect of removal of the pipper from the projected reticle pattern of the Pedestal Sight Manipulation Test. San Antonio, Texas: Human Resources Research Center, Lackland Air Force Base, February 1951. (Research Note 51-2.)

Bilodeau, E. A. & Morin, R. E. Proficiency on the Pedestal Sight Manipulation Test with and without the tracking pipper. San Antonio, Texas: Human Resources Research Center, Lackland Air Force Base, December 1951. (Research Bulletin 51-27.)

Binet, A. L' Etude Experimentale de L' Intelligence. Paris: Costes, 1903.

Blodgett, H. C. The effect of introduction of reward upon the maze performance of rats. Univ. Calif. Publ. Psychol., IV, No. 8, 1929.

Book, W. F. & Norvelle, L. The will to learn: an experimental study of incentives in learning. Pedagog. Sem., 1922, 29, 305-362.

Brown, J. S. A proposed program of research on psychological feedback (knowledge of results) in the performances of psychomotor tasks. In research Planning Conference on perceptual and motor skills. San Antonio, Texas: Human Resources Research Center, Lackland Air Force Base, November 1949. (Conference Report 49-2.) Pp. 81-87.

Bruce, R. H. The effect of removal of reward on the maze performance of rats. Univ. Calif. Publ. Psychol., 1930, IV, No. 13.

Buros, Oscar. The Sixth Mental Measurements Yearbook. New Jersey: The Gryphon Press, 1965.

Crafts, L. W., & Gilbert, R. W. The effect of knowledge of results on maze learning and retention. J. Educ. Psychol., 1935, 26, 177-187.

Crawley, S. L. An experimental investigation of recovery from work. Arch. Psychol., 1926, XIII, No. 85.

Cronback, Lee. Essentials of Psychological Testing. New York: Harper & Bros., 1960.

Dees V., & Grindley, G. C. The effect of knowledge of results on learning and performance: IV. The direction of the error in very simple skills. Quart. J. Exper. Psychol., 1951, 3, 36-42.

DuBois, P. H. A Test-Dominated Society: China, 1115 B.C.-1905 A.D. Proceedings - 1964 Conference-E.T.S., 1965, 3-11.

Eaton, M. T. A study of latent learning. J. Exper. Psychol., 1935, 18, 683-707.

Etwell, J. L., & Grindley, G. C. The effect of knowledge of results on learning and performance: I. A coordinated movement of the two hands. Brit. J. Psychol., 1938-9, 39, 39-54.

Galton, F. Inquiries into Human Faculty and its Development. London: Macmillan, 1883.

Gibson, E. J. & Smith, J. The effect of training in distance on the judgment of size-at-a-distance. San Antonio, Texas: Human Resources Research Center, December 1952. (Research Bulletin 52-39.)

Grindley, G. C. The formation of a simple habit in guinea pigs. Brit. J. Psychol., 1932, XXIII, 127.

Guilford, J. P. Psychometric Methods. New York: McGraw-Hill, 1954.

Guilford, J. P. Fundamental Statistics in Psychology and Education. New York: McGraw-Hill, 1950.

Gulliksen, H. Theory of Mental Tests. New York: Wiley, 1950.

Helmstadter, G. C., & Ellis, D. S. Rate of manipulative learning as a function of goal-setting techniques. J. Exper. Psychol., 1952, 43, 125-129.

Johanson, A. M. Influence of incentive and punishment on reaction time. Arch. of Psychol., 1922, 8, No. 54.

Judd, C. H. Practice without knowledge of results. Psychol. Rev. Monog. Suppl., 1905-6, 7, 185-198.

Keller, F. S. Studies in International Morse Code. I. A new method of teaching code reception. J. Appl. Psychol., 1943, 27, 407-415.

Lorge, I., & Thorndike, E. L. The influence of delay in the after-effect of a connection. J. Exper. Psychol., 1935, 18, 186-194.

MacPherson, S. J., Dees, V., & Grindley, G. C. The effect of knowledge of results on learning and performance: II. Some characteristics of very simple skills. Quart. J. Exper. Psychol., 1948-9, 1, 68-78.

MacPherson, S. J., Dees, V., & Grindley, G. C. The effect of knowledge of results on learning and performance. III. The influence of the time interval between trials. Quart. J. Exper. Psychol., 1948-9, 1, 167-174.

Morgan, C. L., & Morgan, L. V. Effects of immediate awareness of success and failure upon objective examination scores. J. Exper. Educ., 1935, 4, 63-66.

Morin, R. E. The functions of informative feedback and rewarding feedback in acquisition of a lever positioning habit. San Antonio, Texas: Human Resources Research Center, Lackland Air Force Base, December 1951. (Research Note 51-11.)

Morin, R. E., & Gagne, R. M. Pedestal Sight Manipulation Test performance as influenced by variations in type and amount of psychological feedback. San Antonio, Texas: Human Resources Research Center, Lackland Air Force Base, October 1951. (Research Note 51-7.)

Nunnally, J. C. Tests and Measurements. New York: McGraw-Hill, 1959.

Pressey, S. L. Development and appraisal of devices providing immediate automatic scoring of objective tests and concomitant self-instruction. J. of Psychol., 1950, 29, 417-447.

Rethlingshafer, D. Motivation. New York: McGraw-Hill, 1963.

Ross, C. C. An experiment in motivation. J. Educ. Psychol., 1927, 18, 337-346.

Ross, C. C. The influence upon achievement of a knowledge of progress. J. Educ. Psychol., 1933, 24, 609-619.

Saltzman, I. J. Delay of reward and human verbal learning. J. Exper. Psychol., 1951, 41, 437-439.

Seashore, H., & Bavelas, A. The functioning of knowledge of results in Thorndike's line-drawing experiment. Psychol. Rev., 1941, 48, 155-164.

Seashore, R. H., Underwood, B. J., Houston, R., & Berks, L. The influence of knowledge of results on performance. Unpublished manuscript.

Smode, A. F. Learning and Performance in a Tracking Task Under Two Levels of Achievement Information Feedback. J. Exp. Psychol., 1958, 56, 597-704.

Terman, L. W. The Measurement of Intelligence. Boston: Houghton Mifflin Co., 1916.

Terman, L. W., & Merrill, M. Measuring Intelligence. Boston: Houghton Mifflin Co., 1937.

Thorndike, E. L. The law of effect. Amer. J. Psychol., 1927, 39, 212-222.

Thorndike, E. L. Human Learning. New York: Century, 1931.

Thorndike, E. L. The Fundamentals of Learning. New York: Bureau of Publications, 1932, Teachers' College, Columbia Univ.

Thorndike, E. L. The Psychology of Wants, Interests, and Attitudes. New York: Century, 1935.

Trowbridge, M. H., & Cason, H. An experimental study of Thorndike's theory of learning. J. Gen. Psychol., 1932, 7, 245-260.

Waters, R. H. The specificity of knowledge of results and improvement. Psychol. Bull., 1933, 30, 673. (Abstract.)

LAST FIRST MIDDLE

ANENT ADDRESS _____
STREET

CITY STATE

AL SCHOOL _____

OF DENTAL APTITUDE TEST _____
MONTH YEAR

DIRECTIONS:
Erase the block which corresponds to the answer you select for each question.

If the letter "C" appears under the block you erase, this signifies that your choice is CORRECT and you should continue to the next question.

If the letter "I" appears under the block you erase, this signifies that your choice is INCORRECT. If an "I" appears continue to erase your next best choice until the correct answer is determined.

For the purpose of this project, multiple erasures will be scored as errors. DO NOT randomly erase blocks since this will penalize your total score.

QUANT.	No. Right
VERBAL	No. Right
TOTAL	No. Right

1 C I 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39

41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79

81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110

LAST FIRST MIDDLE

NENT ADDRESS _____
STREET

CITY STATE

L SCHOOL _____

OF DENTAL APTITUDE TEST _____
MONTH YEAR

DIRECTIONS:

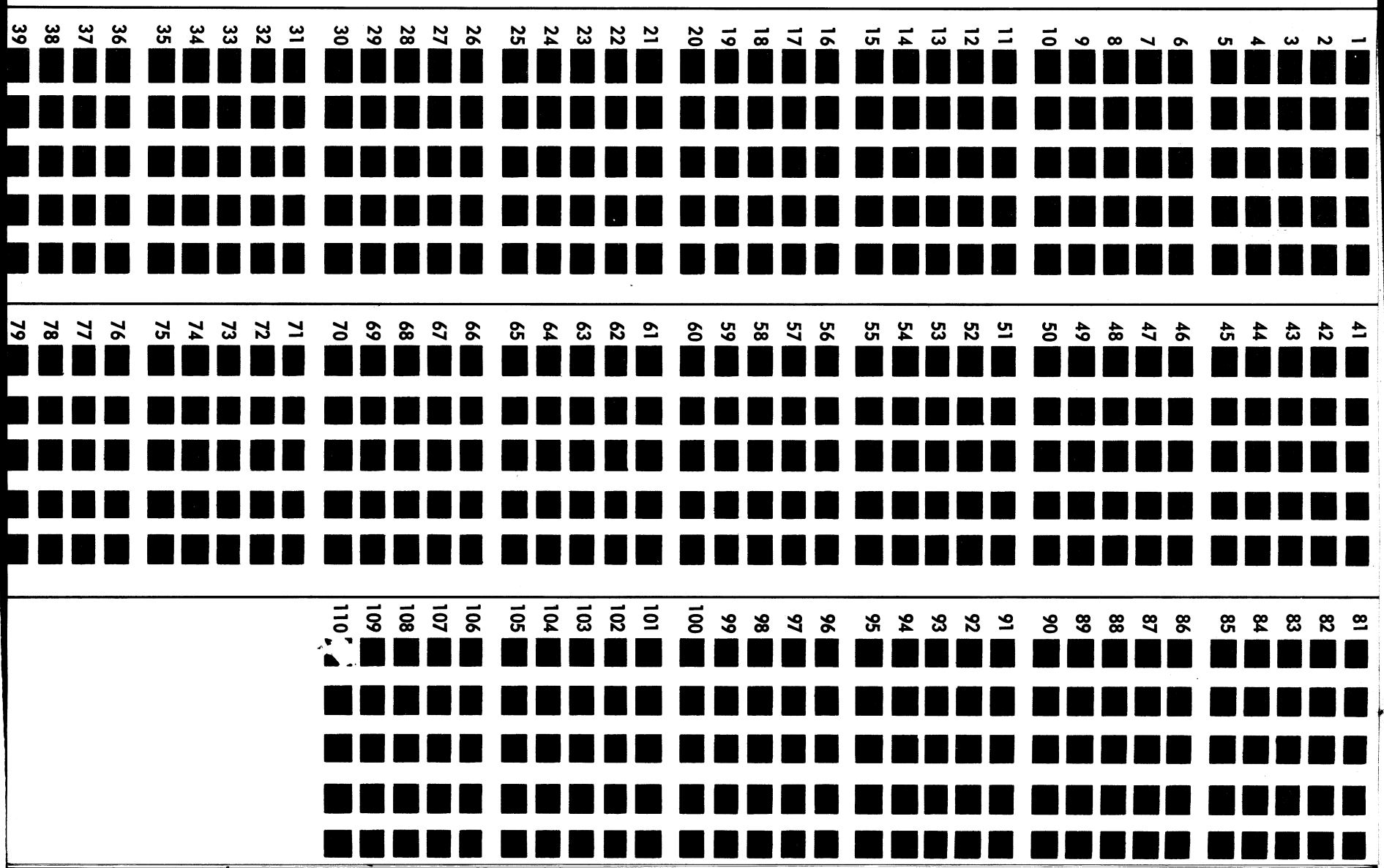
Erase the block which corresponds to the answer you select for each question.

For the purpose of this project, multiple erasures will be scored as errors. DO NOT randomly erase blocks since this will penalize your total score.

QUANT. No. Right

VERBAL No. Right

TOTAL No. Right



APPROVAL SHEET

The dissertation submitted by Thomas John Ginley has been read and approved by a board of three members of the Department of Psychology.

The final copies have been examined by the director of the dissertation and the signature which appears below verifies the fact that any necessary changes have been incorporated, and that the dissertation is now given final approval with reference to content, form, and mechanical accuracy.

The dissertation is therefore accepted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

January 16, 1967
Date

Edmund P. Marx
Signature of Adviser