GREBENKEMPER, NANCY S. The Effect of Social Reinforcement Contingencies on the Predictive Validity of IQ Test Scores. (1976) Directed by: Dr. Marilyn Erickson. Pp. 29.

One important use of intelligence testing is to predict future academic success. The achievement test is one measure of academic success. If intelligence tests are to be useful, they must demonstrate good predictive validity. It was hypothesized that contingent social reinforcement for administration of both tests would maximize predictive validity.

Sixty-four kindergarten and first grade white males were administered the Slosson Intelligence Test and the Reading Test of the Wide Range Achievement Test. Half of the subjects were given the intelligence test under the contingent social reinforcement condition in which praise was presented contingently for correct responses. The remaining half of the subjects were administered the intelligence test under the noncontingent social reinforcement condition in which praise was presented after every other response regardless of accuracy. Half of each of the two groups received contingent social reinforcement and half received noncontingent social reinforcement during the administration of the achievement test. Thus, there were four experimental groups.

The results indicated no significant difference between contingent social reinforcement and noncontingent social reinforcement on both the intelligence and achievement tests. Furthermore, there was no significant difference between the

correlation coefficients for the four experimental conditions indicating that specification of social reinforcement contingency had no effect on predictive validity.

THE EFFECT OF SOCIAL REINFORCEMENT CONTINGENCIES ON THE PREDICTIVE VALIDITY OF IQ TEST SCORES

by

Nancy S. Grebenkemper

A Thesis Submitted to
the Faculty of the Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Master of Arts

Greensboro
1976

Approved by

Marilyn T. Erickson
Thesis Adviser

APPROVAL PAGE

This thesis has been approved by the following committee of the Faculty of the Graduate School at the University of North Carolina at Greensboro.

Thesis Adviser Marilyn T. Erickson

Committee Members D. Evelyn LeSure

Jacquelyn Saelens

June 28, 1976

Date of Acceptance by Committee

ACKNOWLEDGMENTS

I would like to thank Dr. Marilyn Erickson for her constant help and encouragement throughout this endeavor. In addition, I would like to thank Dr. Jacqueline Gaebelein and Dr. Evelyn LeSure for their critical readings of this paper.

Special thanks are due to the principals, teachers, and students of the Reidsville School District for their willing cooperation.

I also gratefully acknowledge the examiners whose cooperation was essential to the completion of the project: Lurlene Alston, Eva B. Carawan, Jamie Carlyle, Melodie Garner, Sue Gizinski, Kathryn Johnson, Lillian Johnson, and Evelyn Walser.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Skinner in Science and Human Behavior (Skinner, 1953) maintained that the best method for the scientific study of behavior is a functional analysis of behavior. Included within this functional analysis are the contingencies of reinforcement. A contingency of reinforcement is defined as the relationship between a behavior and its consequences.

Probability of the response has been demonstrated to be under the control of reinforcement contingencies. Early researchers in experimental psychology found that a response increases in frequency when a positive consequence is made contingent upon the occurrence of that response. Thorndike termed this process the Law of Effect and hypothesized that behavior is stamped in when followed by positive consequences. Although the conceptualization of a stamping-in process is unparsimonious, the idea that a response is more probable after being positively reinforced has been accepted as one of the basic laws of learning.

In order to strengthen the prediction of a response reoccurring given the same stimulus conditions, the contingencies of reinforcement must be specified precisely. The specification of contingencies of reinforcement is reflected in the types of reinforcement schedules: continuous, interval,

and ratio. The random reinforcement of responses is designated as noncontingent reinforcement. Since no contingency of reinforcement is specified, i.e., any response can be followed by reinforcement, the probability of differential increases in a particular response is low. It is only when a specific response is differentially reinforced that the response is expected to increase in frequency.

Most researchers and practitioners in the field of applied psychology have realized the importance of reinforcement contingencies in the functional analysis of behavior. The developers of intelligence tests, however, have not attended to the specification of reinforcement contingencies with regard to task item responses in standardized test procedures. They have specified only the stimulus antecedents for each item and the criteria for correct responses. In addition, test developers have advocated the use of praise for the purpose of motivating the client. However, neither the exact nature of reinforcement nor the reinforcement contingency has been dealt with adequately by intelligence test developers. For example, according to the test manual noncontingent praise is considered appropriate during the administration of the Slosson Intelligence Test. In the WISC-R manual, Wechsler (1974) recommends the use of "judicious reinforcement of effort" to encourage the child. Instructions in the Stanford-Binet manual (Terman & Merrill, 1972) say, "In general, it is effective to praise frequently

and generously, but if this is done in too lavish and stilted a fashion, it is likely to defeat its own purpose." The manual does suggest that praise be withheld until completion of the item.

Jerome Sattler, author of a recent textbook on the assessment of intelligence, specifies that administration procedures have to be flexible in order to elicit good test-taking behavior and to obtain a valid assessment of intellectual functioning (Sattler, 1974). Some clinicians instruct the beginning psychological examiner to give "brief, natural, and casual praise" for success in order to establish rapport (Martin, 1941), while others advise the examiner to praise the child for his effort rather than for the results of his effort (Cole, 1953). Instructions in intelligence test manuals and books on intelligence assessment give no clear indication of when verbal reinforcement is to be used and under what conditions; that is, there are no clear guidelines with respect to reinforcement contingencies.

Beginning in the 1920's a number of investigators have examined the effect of reinforcement on IQ scores. Several of them have found that test scores increased when reinforcement followed correct responses (Ayllon & Kelly, 1972; Clingman & Fowler, 1976; Edlund, 1972; Hurlock, 1924, 1925; Smeets & Striefel, 1975). However, other researchers have found that test scores are not increased under reinforcement conditions (Benton, 1936; Clingman & Fowler, 1975; Maller & Zubin, 1932; Tiber & Kennedy, 1964).

Hurlock was one of the first investigators to examine the effect of manipulation of motivational conditions on IQ scores. Third, fifth, and eighth grade children were first administered either the National Intelligence Test or the Otis Self Administering Test under standard conditions. Matched on age, sex, race, and intelligence test scores, children were divided into three groups: praise, punishment and control. Following the initial testing, groups of children were either praised for their score, given negative verbal feedback for their score, or no comment was made concerning their score. After one week, the children were retested on a different form. Results indicated that praise and reproof were significantly more effective in raising IQ scores compared to the control group. It should be noted that members of each experimental condition were not individually reinforced or punished, and that test scores, not specific responses, were reinforced.

In a second experiment, Hurlock (1925) studied the effect of individual reproof and praise on test scores of the modified Courtis Research Tests in arithmetic over several test administrations. On the basis of the first test, children matched on age, sex, and initial ability were assigned to one of four conditions: praise, reproof, ignored, and control. All groups except the control group were tested in the same room. A different test form was used for each of the five days. Before each test administration, individuals from the praise and reproof conditions were called in front

of the class and either praised and encouraged or given negative verbal feedback for their test performance. Results showed that the test scores of the praised group increased significantly when compared to the scores of the other groups.

Maller and Zubin (1932) followed the procedures of Hurlock's first experiment, but their research differed in two respects: (1) rivalry was substituted for praise and reproof as the motivational factor, and (2) the same form of the National Intelligence Test was used. Thirteen days following the first administration, the children were retested. The experimental group was told that a prize would be awarded to each person who improved his relative standing. Scores for both the experimental and control groups increased but were not significantly different. Further analysis showed that the experimental group attempted more items but were less accurate.

Benton (1936) incorporated Maller and Zubin's conceptualization of rivalry into his motivational condition which included the promise of a prize if the child's relative standing improved and praise from the principal for good work and encouragement to do better. The Otis Self Administering Test was administered to 50 seventh and eighth grade children. Matched on age, test score, sex, and grade, the children were divided into a control and a motivational group. The test was readministered after 28 days. There was no significant difference between groups.

This early research utilized group administered tests to study the effects of different motivational conditions on test performance. Reinforcement or the promise of reinforcement was delivered for total test scores, not for individual responses. Furthermore, in all but Hurlock's second experiment the whole group, not the individual, was reinforced or promised reinforcement.

Later researchers used individually-administered intelligence tests to study the effects of reinforcement contingencies on IQ scores. Tiber and Kennedy (1964) divided 480 second and third grade children selected equally from three social groups (middle-class white, lower-class white, lower-class black) into four groups: verbal praise, verbal reproof, candy reward, and control. Each subject was administered the Stanford-Binet Intelligence Scale, Form L-M, with the appropriate incentives administered at the end of each item. The rewards, however, were not contingent on correct responses. Results indicated no significant difference between groups.

More recent research, with the exception of Clingman and Fowler (1975), has shown better test performance when correct responses were reinforced. Ayllon and Kelly (1972) studied the relative effect of reinforcement on test performance with both retarded and normal fourth-grade children. Using a test-retest procedure, trainable retardates were administered the Metropolitan Reading Test under two conditions: a standard testing condition and token reinforcement condition. The standard testing condition was not described.

The token reinforcement condition consisted of the child receiving one token for each correct response at the completion of each subtest. Using a normal population, the above procedures were replicated except that two forms of the Metropolitan Achievement Test were used as the dependent measures. In both experiments, test scores under the token reinforcement administration were significantly higher than under the standard test procedures. The first experiment with retardates suffered from the lack of a group to control for practice effects. Furthermore, in both experiments, reinforcement was delivered on a delayed schedule.

Edlund (1972) studied the effect of reinforcement contingent on correct responses on Stanford-Binet IQ scores. Twenty-two 5-7 year-old children from low-middle and lower class homes were given Form L of the revised Stanford-Binet under standard testing conditions. These children were then divided into two groups matched on IQ score, age, sex, and liking of candy. Both the control and experimental groups were administered Form M of the Stanford-Binet seven weeks following the first test administration. For the control group, standard testing procedures were followed. For the experimental group, each child received a M&M candy following each correct response. IQ scores of children in the experimental group increased significantly when compared with the IQ scores of children in the control group.

Clingman and Fowler (1975) attempted to replicate Edlund's experiment with several modifications. Subjects were

slightly older, from a higher SES class, and initially had considerably higher IQ scores than the subjects of Edlund's study. In addition, a third condition was added to study the effects of noncontingent reinforcement. Following administration of the Stanford-Binet, Form L, to first and second grade children of above-average intelligence, the children were randomly assigned to one of three groups: (1) contingent reinforcement where the child was given candy following each correct response, (2) yoked-comparison where the child received the same amount of candy as a child in the contingent reinforcement condition but reinforcement preceded a response, and (3) a no-candy control. Six weeks later the children were retested on Form M of the Stanford-Binet under the different reinforcement conditions. Results revealed no statistical significant difference between groups. Clingman and Fowler offered several explanations to account for their results which are contradictory to those of Edlund: above-average intelligence of Clingman and Fowler's subjects and lower motivational levels of Edlund's children. The children in Clingman and Fowler's experiment were hypothesized to have been performing at an optimal level as reflected by the high IQ scores; thus, reinforcement had no effect. Clingman and Fowler suggested, in addition, that Edlund's results could be explained as a function of regression to the mean. They also criticized Edlund's study for a selection bias.

In a second study, Clingman and Fowler (1976) examined the hypothesis that a motivational deficit in low IQ children accounted for the discrepancy between Clingman and Fowler's earlier study and Edlund's study. Following a standard administration of Form A of the Peabody Picture Vocabulary Test, 72 first and second grade children were divided on the basis of IQ scores into three blocks: low, middle, and high. From each block, subjects were randomly assigned to one of three conditions: contingent candy reward, noncontingent candy reward, or no reward. Four weeks later the children were retested on Form B of the Peabody Picture Vocabulary Test under the different reinforcement conditions. The results showed that candy reinforcement contingent on correct responses significantly increased IQ scores of children from the low IQ group but had no effect on IQ scores of children from the middle and high IQ groups.

Smeets and Striefel (1975) compared the effects of reinforcement on eleven to eighteen year-old multihandicapped deaf children's performance on the Raven Progressive Matrices. Using a test-retest paradigm, children were administered the Raven Progressive Matrices under standard testing conditions receiving their choice of candy or pennies at the end of the test. Children were then randomly assigned to one of four groups: (1) end of session reinforcement, (2) noncontingent reinforcement where every response was reinforced, (3) delayed reinforcement where reinforcement was delivered for correct

responses after a set number of items, and (4) immediate contingent reinforcement. Reinforcement consisted of checkmarks which could be traded for pennies or candy. The children were retested and received checkmarks depending on their respective experimental condition. The immediate contingent reinforcement condition yielded a significantly higher increase on test scores than the other conditions. One fault of this study was that the amount of reinforcement was not controlled.

Results of studies reviewed above have been inconclusive as to the effect of reinforcement on standardized test scores. Individual procedures and different populations may account for the discrepant findings. However, studies that show differential effects due to contingent reinforcement must be assessed with respect to their implications for interpretation of test scores. The primary purpose of intelligence tests has been to aid school personnel in making decisions regarding appropriate academic placement of children. Edlund (1972) stressed the importance of precise reinforcement procedures in obtaining an accurate statement of an individual's academic achievement. Since Edlund views the tester-testee relationship as a component of the IQ score, this relationship must be standardized in terms of reinforcement conditions in order to procure an accurate IQ score. Ayllon and Kelly (1972) recommended the use of contingent reinforcement for correct responses to obtain optimal performance which is a representative reflection of a child's academic performance.

Smeets and Striefel (1975) maintained that reinforcement contingent on correct responses constitutes optimal motivational conditions in obtaining an accurate and valid reflection of the child's abilities. Reviewing the research of Ayllon and Kelly and Edlund, O'Connor and Weiss (1974) have stated that the application of contingent reinforcement during standardized testing can eliminate or account for a portion of the error variance. This reduction in the standard error of measurement would increase the test's reliability and predictive validity.

Since the primary purpose of intelligence tests has been to predict academic achievement, it is important to study the effect of reinforcement contingencies on the predictive validity of IQ scores. There are several ways to conceptualize predictive validity of test results. Correlations between IQ tests and achievement tests, however, provide the most accurate measure due to the fact that they are standardized assessment tools, indicating an established reliability and validity. Volume VII of the Buros' Mental Measurements Yearbook cited three studies investigating the predictive validity of Stanford-Binet IQ scores with various achievement tests (Wide Range Achievement Test, California Achievement Test, Iowa Test of Basic Skills). Results of these studies show correlation coefficients ranging from .40 to .77 for individual subtests and correlation coefficients from .59 to .74 for composite scores (Churchill & Smith, 1966; Hirshoren, 1969; Washington & Teska, 1970). These correlations

between IQ test scores and achievement test scores are less than optimal. Standardization of reinforcement contingencies in test administration could eliminate a portion of the error variance in the test scores, thereby increasing the correlation between achievement and IQ tests, and, in effect, increasing the predictive validity of intelligence tests.

The purpose of the present study was to investigate the effects of social reinforcement contingencies on standardized test scores. This experiment compared the effect of contingent social reinforcement and noncontingent social reinforcement on IQ and achievement test scores in order to determine the contingency arrangement that maximized predictive validity of intelligence tests.

Although other investigators have used tangible reinforcers contingent on correct responses as a means to motivate the child and procure a more accurate assessment of the child's skills, this study used praise as a reinforcer. Test manuals indicate that the use of tangible reinforcers is unacceptable in a standardized testing situation, and that praise is to be used to motivate the child. Thus, in order to parallel the real testing situation more closely, praise presented either contingently or noncontingently was used in the testing situation.

It was predicted that standardized test scores would be higher for children who received social reinforcement contingent on correct responses than for children who received noncontingent social reinforcement. Furthermore, the group

who received contingent reinforcement for administration of both intelligence and achievement tests would show the highest correlation coefficient, while the group that received noncontingent reinforcement for both test administrations would show the lowest correlation coefficient.

CHAPTER II

METHOD

Subjects

Subjects were thirty-two white male kindergarten and thirty-two white male first grade children (mean CA = 6.57 years) attending public elementary schools in the Reidsville, North Carolina, school district.

Examiners and Dependent Variables

Two dependent variables, the Slosson Intelligence Test for Children and Adults (Slosson, 1963) and the Reading Test of the Wide Range Achievement Test (Jastak & Jastak, 1965) were used to assess the effect of the different reinforcement contingencies. Both instruments have proven to be both reliable and valid. Reliability and validity coefficients for the Slosson range in the .90's as reported in the Slosson manual. The WRAT manual indicates for the Reading Test of the WRAT a split-half reliability coefficient of .987; correlations between the Reading Test of the WRAT and the Verbal Scale of the Wechsler Intelligence Scale for Children range from .68 to .78.

Eight female undergraduate students served as examiners. A graduate student, familiar with the testing instruments, trained the examiners on test procedures. This training consisted of practice administering the tests including practice

in giving the required feedback for all conditions. Test administration was monitored by the graduate student.

Experimental Design

The basic experimental design was a one-way completely randomized design with contingency (contingent vs. noncontingent) as the independent variable.

Independent Variable

The independent variable that was manipulated was the reinforcement contingency. For the administration of the IQ test and the achievement test, the two groups of thirty-two children each were randomly assigned to one of four conditions. For the administration of the intelligence test, half of the children received contingent reinforcement; the remaining half received noncontingent reinforcement. Half of each of the above two groups received contingent reinforcement, and half received noncontingent reinforcement for the administration of the achievement test.

Contingent reinforcement consisted of every correct response being followed by social reinforcement, i.e., "Very good," "You're doing fine," from a prescribed list, delivered with expression and a smile. For the noncontingent reinforcement condition, social reinforcement was given for every other response regardless of whether the response was correct or not. The tests are designed such that, on the average, fifty percent of the responses are correct, and fifty percent of the responses are incorrect; thus, the amount of

reinforcement is comparable in the contingent and noncontingent conditions. Thus, the four experimental conditions were: contingent-contingent, contingent-noncontingent, noncontingent-contingent, noncontingent-noncontingent.

Procedure

Subjects were individually administered the Slosson Intelligence Test and two weeks later the Reading Test of the Wide Range Achievement Test by the same examiner. Subjects were randomly assigned to examiners with the restriction that two subjects from each condition be assigned to each of the eight examiners.

Statistical Analysis

Since the only examiners available for the study were four black and four white undergraduates, the statistical analysis includes race of examiner as a factor. Two two-way analyses of variance were performed to determine the effect of the different contingencies and the race of the examiner on both the IQ and the achievement scores. In addition, correlation coefficients between the IQ and achievement test scores were computed to determine the effect different social reinforcement contingencies have on the predictive validity of IQ scores. The correlations were tested for statistical significance and compared by statistical analysis.

CHAPTER III

RESULTS

Since half of the examiners were black and half white, a two-way analysis of variance with race (black vs. white) and contingency (contingent vs. noncontingent) as factors was performed on both the Slosson IQ scores and the standard scores on the Reading Test of the WRAT. No significant differences were obtained on either of the factors for both analyses. Table 1 and 2 present summaries of the statistical analyses.

Table 3 presents the means and standard deviations of the Slosson IQ scores and the WRAT Reading Test standard scores for the two experimental conditions: praise presented contingently for correct responses and praise presented for every other response regardless of correctness. On the Slosson Intelligence Test, when praise was contingent on correct responses, the mean IQ score was 99.06 with a standard deviation of 14.16 points; whereas, when praise was noncontingent, the mean IQ score was 100.09 with a standard deviation of 15.27 points. On the Reading Test of the WRAT, when praise was delivered contingently, the mean standard score was 97.06 with a standard deviation of 15.79 points; when praise was delivered noncontingently, the mean standard score was 96.31 with a standard deviation of 8.78 points.

Table 1

ANOVA on IQ Scores

Source	SS	df	MS	F	p
Race	153.141	1	153.141	.6755	n.s.
Contingency	17.016	1	17.016	.0751	n.s.
Race x Contingency	129.391	1	129.391	.5708	n.s.
Error	13601.860	60	226.698		

Table 2

ANOVA on Achievement Scores

Source	SS	df	MS	F	p
Race	400.000	1	400.000	2.4789	n.s.
Contingency	9.000	1	9.000	.0558	n.s.
Race x Contingency	361.000	1	361.000	2.2372	n.s.
Error	9681.602	60	161.360		

Table 3

Means and Standard Deviations of IQ and Achievement Test Scores

	Mean Slosson IQ Score	Standard Deviation	Mean WRAT Standard Score	Standard Deviation
Contingent Praise	99.0625	14.1641	97.0625	15.7875
Noncontingent Praise	100.0938	15.2729	96.3125	8.7801

Table 4 presents the Pearson product-moment correlation coefficients between the Slosson IQ scores under the two contingency conditions and the WRAT standard scores under the two contingency conditions. The highest correlation was obtained when the subject received noncontingent praise on the IQ test and praise contingent on correct responses on the achievement test. The lowest correlation was obtained when the subject received contingent praise on the IQ test but noncontingent praise on the achievement test. When the subject was administered both standardized tests under the same contingency conditions (contingent praise or noncontingent praise), the resulting correlations fell between the highest and the lowest correlations. All correlations were significantly different from zero ($p < .01$) but were not significantly different from each other.

Table 4

Correlation Coefficients between IQ and Achievement Scores

		Achievement	
		Contingent Praise	Noncontingent Praise
IQ	Contingent Praise	.72	.57
	Noncontingent Praise	.76	.65

CHAPTER IV

DISCUSSION

The present findings indicate that there was no difference between contingent praise and noncontingent praise on standardized test scores. The data also show that the race of the examiner did not significantly affect standardized test scores.

Recent studies (Allyon & Kelly, 1972; Clingman & Fowler, 1976; Edlund, 1972; Smeets & Striefel, 1975) have shown that contingent reinforcement with candy as the reinforcer has been effective in raising IQ scores. In this study praise was used as a reinforcer. However, the effectiveness of praise as a reinforcer was not empirically demonstrated for this sample of children; other studies have demonstrated the effectiveness of praise as a reinforcer with similar samples of children but not in a standardized testing situation. It may be that praise was not a reinforcer and had no effect, presented either contingently or noncontingently, on test performance. Furthermore, since the standardized tests used only required a short time for administration, it may have been difficult to develop a relationship between the examiner and the subject. Thus, the conditions may not have been sufficient for praise to function as a reinforcer.

Clingman and Fowler's (1976) recent study revealed an interaction effect between initial level of IQ and

reinforcement contingency. They found that candy presented contingently for correct responses increased IQ scores from the low IQ group but had no effect on scores of children from the middle and high IQ groups. The lack of significant findings in the present study might be explained by the fact that the population sampled was normally distributed. Thus, only a small portion was from the lower IQ score group. According to Clingman and Fowler's data, only scores of the children in this group would increase due to contingent rewards. If this is the case, there would be too small a portion of the sample affected by contingent rewards to create a difference between contingent and noncontingent groups.

The correlations between IQ scores and achievement test scores under the different contingency conditions were not significantly different. The data do not support the hypothesis that contingent praise for both tests would result in the highest correlation coefficient.

However, it is interesting to note that the two highest correlations occurred when praise contingent on correct responses was presented during the achievement test regardless of the contingency condition in effect during the IQ testing. On the achievement test the contingent praise condition accounted for 55 percent of the variance, while the noncontingent praise condition accounted for only 37 percent of the variance.

This difference may indicate that contingent praise is more effective in increasing predictive validity when

administering an achievement test compared to an IQ test. Thus, the nature of the test may be an important variable to consider when investigating the effect of contingent versus noncontingent praise on test scores, and, in particular, on predictive validity. The IQ test and achievement test differ in one important respect: homogeneity. The IQ test consists of questions covering a wide spectrum of information while the achievement test includes only a small aspect of a particular skill. The child may be better able to self-monitor his correctness on the achievement test where the distinction between right and wrong responses is more explicit than on the IQ test where there is a larger set of questions which the child is unsure about in respect to the correctness of his response. Further research is needed to examine whether there is an interaction effect between the nature of the test and reinforcement contingencies.

Another explanation for the fact that the contingent praise condition on the achievement test accounted for a greater portion of the variance may be that a positive relationship between the examiner and the subject was developing by the time of the second testing. That is, the child needed a certain period of time to observe that the new adult's behaviors matched those of other reinforcing adults; this exposure allowed the new adult to acquire reinforcing properties. Thus, the total amount of time spent contingently reinforcing the subject may, to a large extent, determine the reinforcing effectiveness of the examiner.

Another factor to take into consideration is the possibility of an interaction effect between the sex of the examiner and the sex of the subject; that is, a study including male examiners and female subjects may generate different findings.

Future research in this area should involve the study of reinforcement contingencies incorporating the modifications as outlined above. It would be particularly informative to replicate Clingman and Fowler's (1976) recent study using praise as a reinforcer. Research should also continue to examine variables that affect test scores in an effort to improve the quality of standardization and especially to increase the validity of these instruments.

BIBLIOGRAPHY

Ayllon, T., & Kelly, K. Effects of reinforcement on standardized test performance. Journal of Applied Behavior Analysis, 1972, 5, 477-484.

Benton, A. L. Influence of incentives upon intelligence test scores of school children. Journal of Genetic Psychology, 1936, 49, 494-497.

Buros, O. K. (Ed.). The Seventh Mental Measurements Yearbook. Highland Park, N. Y.: Gryphon Press, 1972.

Churchill, W. D., & Smith, S. E. The relationship of the 1960 Revised Stanford-Binet Intellignece Scale to achievement test scores over a three year period. Educational and Psychological Measurement, 1966, 26, 1015-1020.

Clingman, J., & Fowler, R. L. The effects of contingent and noncontingent rewards on the IQ scores of children of above-average intelligence. Journal of Applied Behavior Analysis, 1975, 8, 90.

Clingman, J., & Fowler, R. L. The effects of primary reward on the I.Q. performance of grade-school children as a function of initial I.Q. level. Journal of Applied Behavior Analysis, 1976, 9, 19-23.

Cole, D. Communication and rapport in clinical testing. Journal of Consulting Psychology, 1953, 17, 132-134.

Edlund, C. V. The effect on the behavior of children, as reflected in the IQ scores, when reinforced after each correct response. Journal of Applied Behavior Analysis, 1972, 5, 317-319.

Hirshoren, A. A comparison of the predictive validity of the revised Stanford-Binet Intelligence Scale and the Illinois Test of Psycholinguistic Abilities. Exceptional Children, 1969, 70, 517-521.

Hurlock, E. B. The value of praise and reproof as incentives for children. Archives of Psychology, 1924, 11, No. 71, 78.

Hurlock, E. B. An evaluation of certain incentives used in schoolwork. Journal of Educational Psychology, 1925, 16, 145-159.

Jastak, J. F., & Jastak, S. R. The Wide Range Achievement Test. New York: Psychological Corporation, 1965.

Maller, J. B., & Zubin, J. The effects of motivation upon intelligence test scores. Journal of Genetic Psychology, 1932, 41, 136-151.

Martin, M. F. Gaining rapport in the school clinic. Mental Hygiene, 1941, 25, 251-255.

O'Connor, J. J., & Weiss, F. L. A brief discussion of the efficacy of raising standardized test scores by contingent reinforcement. Journal of Applied Behavior Analysis, 1974, 7, 351-352.

Sattler, J. M. Assessment of children's intelligence. Philadelphia, Pa.: W. B. Saunders Co., 1974.

Skinner, B. F. Science and human behavior. New York: MacMillan Co., 1953.

Slosson, R. L. Slosson Intelligence Test (SIT) for children and adults. New York: Slosson Educational Publications, 1963.

Smeets, P. M., & Striefel, S. The effects of different reinforcement conditions on the test performance of multihandicapped deaf children. Journal of Applied Behavior Analysis, 1975, 8, 83-89.

Terman, L. M., & Merrill, M. A. Stanford-Binet Intelligence Scale. Boston: Houghton Mifflin, 1972.

Tiber, N., & Kennedy, W. A. The effects of incentives on the intelligence test performance of different social groups. Journal of Consulting Psychology, 1964, 28, 187.

Washington, E. D., & Teska, J. A. Correlations between the Wide Range Achievement Test, the California Achievement Tests, the Stanford-Binet, and the Illinois Test of Psycholinguistic Abilities. Psychological Reports, 1970, 26, 291-294.

Wechsler, D. Manual for the Wechsler Intelligence Scale for Children - Revised. New York: Psychological Corporation, 1974.

APPENDIX A

List of phrases used as Social Reinforcers:

You're good at this

Fine

You're doing well

Very good

You really know how to do this

You're doing a good job

Good

You're doing fine

That's good