The Effects of Monetary Incentives on Group Intelligence Test Performance of Lower Class Children

Larry Gerbig
Western Kentucky University

Follow this and additional works at: https://digitalcommons.wku.edu/theses
Part of the Child Psychology Commons

Recommended Citation
https://digitalcommons.wku.edu/theses/2400

This Thesis is brought to you for free and open access by TopSCHOLAR®. It has been accepted for inclusion in Masters Theses & Specialist Projects by an authorized administrator of TopSCHOLAR®. For more information, please contact topscholar@wku.edu.
Gerbig,
Larry A.
1973
THE EFFECTS OF MONETARY INCENTIVES ON GROUP INTELLIGENCE
TEST PERFORMANCE OF LOWER CLASS CHILDREN

A Thesis
Presented to
the Faculty of the Department of Psychology
Western Kentucky University
Bowling Green, Kentucky

In Partial Fulfillment
of the Requirements for the Degree
Master of Arts

by
Larry A. Gerbig

June 1973
THE EFFECTS OF MONETARY INCENTIVES ON GROUP INTELLIGENCE
TEST PERFORMANCE OF LOWER CLASS CHILDREN

David A. Smith
Director of Thesis

Carl B. Muray

Julia C. Rice

Approved: July 20, 1973
Date

Edward S.
Dean of the Graduate College

Approved: August 8, 1973
Date
ACKNOWLEDGMENTS

I would like to express my sincere appreciation to Dr. D. A. Shiek, my Thesis Committee Chairman, for his patience, generosity, and expertise given during this study. I would like to thank Dr. Sandra Reese and Dr. Carl Martray for their constant help and support. And a special thanks to my wife Karen, for her continued patience and understanding throughout the entire study.
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II. REVIEW OF LITERATURE</td>
<td>3</td>
</tr>
<tr>
<td>III. STATEMENT OF PROBLEM</td>
<td>12</td>
</tr>
<tr>
<td>IV. METHOD</td>
<td>14</td>
</tr>
<tr>
<td>Subjects</td>
<td>14</td>
</tr>
<tr>
<td>Instrument</td>
<td>14</td>
</tr>
<tr>
<td>Design</td>
<td>15</td>
</tr>
<tr>
<td>Procedure</td>
<td>16</td>
</tr>
<tr>
<td>V. RESULTS</td>
<td>18</td>
</tr>
<tr>
<td>VI. DISCUSSION</td>
<td>19</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>23</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

Since the advent of the first intelligence test in 1905, the primary goal of the psychometric field has been to obtain the most reliable and valid evaluation of an individual's level of intellectual functioning. The major difficulty concerning this goal, however, has been how to motivate the individual to work at his optimum level of performance. Terman (1916) attempted to solve this motivation problem through the use of praise. "Exclamations like 'Fine!,' 'Splendid!,' etc. should be used lavishly. Almost any innocent deception is permissible which keeps the child interested, confident, and at his best level of effort [p. 125]." Terman and Merrill (1937) instructed examiners to enlist the individual's best efforts through the establishment of good rapport, or the resulting score would be an underestimate to some unknown degree. In their latest revision, Terman and Merrill (1960) again stressed the importance of rapport, but felt that praise should be given not only for success, but for general effort as well.

Because the individual's best possible performance is sometimes difficult, if not impossible to measure, it is all too often taken for granted by the examiner that the individual's best efforts have been demonstrated without considering the actual amount of motivation and the incentives utilized in the testing environment. If it can be
demonstrated that the use of incentives to maximize the level of motivation is an important factor in the optimal performance of the individual, then the need to control this factor in order to obtain valid measures of intellectual functioning becomes of critical importance.
CHAPTER II

REVIEW OF LITERATURE

In general, studies concerning the effects of incentives upon children's functioning have resulted in some very definite trends. When the practice variable is controlled, praise and material items, such as candy or promise of a prize, have been demonstrated as reasonably stable incentives. They generally have contributed an increase in the functioning level and learning of school children from study to study. Blame, on the other hand, has generally resulted in a decrease in the performance of school children.

Chapman and Feder (1917), using stars and nominal prizes as incentives, found that these incentives exerted a considerable effect on addition, cancellation, and digit-symbol tests, and Hurlock (1925b) found that regardless of age, sex, initial ability, or accuracy, praise was more effective than reproof in improving performance. In a replication of Hurlock's (1925b) study, Cohen (1927) found less difference between the praise and reproof groups, but the same general trends were apparent, while Warden and Cohen (1931) concluded that the commonly used incentives of praise and reproof were no more effective than a simple change in the class routine.

Gilchrist (1916) employed the incentives of praise and reproof on 50 college students taking the Curtiss English Test, and concluded that
encouragement produced more improvement than discouragement, and that
discouragement produced better results than repetition alone.
Silverman (1957) concluded that teacher use of praise and reproof could
not predict the reading growth of students, but could predict the total
verbal output of the children.

Knight and Remmers (1923) found that increased levels of motivation
not only offset extreme fatigue, but produced twice as much work per
unit of time with equal accuracy. Anderson and Smith (1933) supported
these findings in their study. Brenner (1934), on the other hand, found
no difference in the effects of immediate versus delayed praise and
blame on performance.

Blankenship and Humes (1938) attempted to determine the effects of
praise and punishment on the discrimination task performance of 130 Ss,
but found no significant difference between performance during praise
and performance during punishment. Potter (1943) concluded that reproof
had little effect on higher age levels, but varied considerably for the
lower age levels. Terrell and Kennedy (1957), using candy, praise,
reproof, and control motivation on the transposition learning of
children, discovered that the candy group learned the task and trans-
posed this learning significantly more effectively than the other groups,
but that neither praise nor reproof was significantly more effective
than the control conditions.

Other studies have attempted to extend Hurlock's original investi-
gation to encompass the effects of a variety of incentives on the
intelligence test performance of children. These studies, too, have
resulted in some very definite trends. Praise, in general, has been
found to increase intelligence test performance better than reproof or practice alone, while blame has been found to have a debilitating effect on performance.

Hurlock (1924), following the methodology of Gilchrist (1916) and Gates and Rissland (1923), was the first to study the effects of verbal incentives on the intelligence test performance of children. She divided 408 third-, fifth-, and eighth-grade children into groups according to age, sex, race, and intelligence, and administered two forms of the National Intelligence Test to the older children and the Otis SA to the younger children, with a one week interval between testing periods. Praise, reproof, and control groups were used as treatment conditions during the second administration of the test. She concluded that praise was neither superior nor inferior to reproof in increasing intelligence test performance, but that both served to increase performance better than practice alone. In a follow-up study, Hurlock (1925a) replicated her previous investigation and once again found no difference between reproof and praise in increasing IQ scores, but greater performance under these conditions than under practice alone.

Benton (1936) studied not only praise, but strong encouragement, knowledge of results, and promise of a prize as effective incentives in increasing intellectual performance. The children of the incentive group were told their relative standings on the test and were promised a prize if they improved their standing on the second administration of the Otis SA test. In addition, the school principal praised them for their work and strongly urged them to do better on the second test. The
results, however, showed no significant difference between the control group and the incentive group on the second administration of the Otis.

Kennedy, Turner, and Lindner (1962) attempted to study the effectiveness of praise and reproof as a function of level of intelligence of the child. Using two adolescent groups, one group with IQ's from 124-150 and the other group with IQ's from 95-116, they failed to find significant differences in performance of the higher intellect group under all three conditions, but did detect a significant deficit in performance for the average group under the blame incentive. Wilcutt and Kennedy (1963) attempted to test the findings of Kennedy, Turner, and Lindner (1962). Using 90 fourth-grade children, aged 9-11, the children were divided into three IQ groups: 71-90, 91-110, and 111-130. Each of these groups was then subdivided into three experimental groups, praise, blame, or no incentive (control), of five males and five females each. Although no significant relationship between intelligence and verbal incentive effectiveness was revealed, there was a significant increase in performance under praise.

Feldman and Sullivan (1971) studied the effects of rapport enhancement on the intellectual performance of elementary school children. Using 72 students, matched by grade, sex, and Otis IQ, the Ss were divided into enhanced and neutral rapport treatment groups and administered a short form of the Wechsler Intelligence Scale for Children (WISC). The rapport conditions differed in the amount of friendly conversation prior to and during the WISC testing and the inclusion of verbal reinforcement for the first correct response on each WISC subtest. Those children tested under enhanced rapport showed significant
improvement in IQ scores over those tested under the neutral rapport conditions.

Witmer, Bornstein, and Dunham (1971) investigated the effects of three modes of test administration upon the performance of third- and fourth-grade students on four subtests of the WISC, utilizing 90 Ss, randomly assigned to an approval group, a disapproval group, or a control group. Two examiners were then randomly assigned to each of the treatment groups. The examiners gave verbal approval after the first response in each subtest and between subtests to the approval group, verbal disapproval in the same sequence to the disapproval group, and the standard test administration to the control group. For both examiners, the approval group performed significantly higher than the disapproval group. The control group performed higher than the disapproval group and lower than the approval group, although the differences were not significant.

Galdieri, Barcikowski, and Witmer (1972) studied the effects of verbal approval and disapproval on the performance of Middle Class and Lower Class children on the Wechsler Intelligence Scale for Children (WISC). Using 72 Ss from three schools and two examiners randomly assigned to each school, the WISC was administered to each student under both treatment conditions. The verbal approval group was given verbal support for the first response in each subtest. For the first incorrect response in each subtest, they were told, "That was hard, wasn't it? But you're doing good [p. 405]." The disapproval group was given verbal disapproval in the same sequence. The control group was given the test under standardized instructions. Although the Middle Class children had
significantly higher IQ scores than the Lower Class children, no difference between performance under approval and performance under disapproval was evident.

The most recent area of interest has become the investigation of the effects of various incentives on the intellectual performance of different social class and ethnic groups. The lack of research in this area, however, has resulted in a few general trends that are still far from conclusive. Middle and Upper Class children, it seems, appear to perform no better under external incentive conditions of money and food than under the more conventional incentive conditions of praise and grades used in the classroom. Lower Class children, on the other hand, appear to perform better when promised external incentives (food and money) instead of the conventional classroom incentives of praise and grades.

Klugman (1944) sought to determine whether money, rather than praise, would elicit higher performance on the Revised Stanford-Binet Test, 1937 edition. Using 38 white and 34 Negro, Middle Class children, aged 7-14, Klugman employed Form L of the Stanford-Binet. One week later he gave the same students Form M of the instrument. Half of the students were given between five and fifteen cents as a money incentive, depending on the number of correct responses, while the other half were given verbal praise. He found that Negro children performed at a higher level, although not significantly, under the money incentive than under praise. In evaluating this study, however, consideration should be given to the fact that Klugman did not use a control group to determine to what extent incentives actually effected performance.
Tiber and Kennedy (1964) studied the effects of praise, reproof, and candy on the intellectual performance of Middle Class whites, Lower Class whites, and Lower Class Negroes. The students were selected equally from each of the three social groups and randomly assigned to verbal praise, verbal reproof, candy, and control groups. The experimental incentives were administered at the end of each task on the Stanford-Binet Intelligence Scale, Form L-M. Although no difference between incentive groups and no significant interaction between type of incentive and social group was detected, there was a significant difference in IQ scores among these social groups.

Fast (1967), studying the effects of monetary and verbal incentives on the IQ scores of Upper Middle Class and Lower Class children, administered the Wechsler Intelligence Scale for Children (WISC) to 60 white, male, fifth- and sixth-graders. Three months later, the children were divided into control, verbal reward, and monetary reward subgroups and given the WISC again. Rewards consisted of the words "right" or "correct" and a penny for a correct response. No significant difference in performance under either incentive was found. Higgins and Archer (1968) studied the effects of differential rewards on the performance of different social classes on the group administered IPAT Culture Fair Intelligence Test. Using extrinsic rewards (food and money) and intrinsic rewards (praise and grades), the 250 Ss were randomly assigned to one of the treatment conditions, then retested on this instrument. Those Lower Class students who received extrinsic rewards performed significantly greater than those Lower Class students receiving intrinsic rewards. The Upper Class students who received extrinsic rewards,
however, performed no differently than Upper Class students who received the conventional rewards of praise and grades.

In summary, the use of incentives in improving classroom achievement, task performance, and intelligence test performance resulted in some definite trends. Praise and material incentives were demonstrated as reasonably stable incentives in improving classroom achievement and task performance, while blame fairly consistently had a debilitating effect on performance. Praise also had improved intelligence test performance better than reproof or practice alone, while blame had inhibited performance.

The study of the effects of various incentives on the intelligence test performance of different social class and ethnic groups, was the most promising, yet the least investigated topic. For Middle and Upper Class children, the type of incentive used to increase intelligence test performance seemed to have no effect. Children in these categories were motivated by the presence of the incentive itself, regardless of the type. Lower Class children, in contrast, differed in performance according to the type of incentive being offered. Since prior studies yielded inconclusive results, there was a need for further study in this area to determine what incentives, if any, would significantly improve the intelligence test performance of Lower Class children.

The use of monetary incentives in improving the intelligence test performance of Lower Class children has not produced very promising results, but past investigators only studied the effects of monetary incentives in relation to the effects of other intrinsic incentives, such as verbal praise and grades. The differential effects of various
magnitudes of incentives has been neglected. Therefore, studying intelligence test performance of Lower Class children as simply a function of the magnitude of the monetary incentive, rather than a comparison with other incentives, was appropriate.
CHAPTER III

STATEMENT OF PROBLEM

The present study attempted to determine if monetary incentives would produce an effect on intelligence test functioning for Lower Class children. More specifically, the purpose was to determine if Lower Class Ss would obtain a higher intelligence quotient if monetary incentives were employed, and if various amounts of money would have differential effects on intelligence test performance of Lower Class children. In addition, the effects of the incentives when the Ss were divided as to intelligence quotient and sex was investigated. Therefore, the following null hypotheses were tested:

1. There is no significant difference between the change in intelligence quotients of the high IQ group and the change in intelligence quotients of the low IQ group.

2. There is no significant difference among the change in intelligence quotients of the control group and changes in intelligence quotients of the groups receiving incentives of ten cents, twenty-five cents, and fifty cents.

3. There is no significant difference between the change in intelligence quotients of the male group and the change in intelligence quotients of the female group.
4. There is no significant interaction between the amount of incentive and IQ level.

5. There is no significant interaction between sex and IQ level.

6. There is no significant interaction between amount of incentive and sex.

7. There is no significant interaction among IQ level, incentive, and sex.
CHAPTER IV

METHOD

Subjects

The Ss were 101 third-, fourth-, fifth-, and sixth-grade students, 46 of whom were male and 55 female. Of the 101 students, 89 were Negro and the remaining 12 were white. The children in the study were considered to be primarily Lower Class since the school had been allocated Title I funds according to the guidelines of the Elementary and Secondary Education Act of 1965.

Instrument

The instrument used in this study was the Henmon-Nelson Test of Mental Ability, Form A, a group administered intelligence test with a 30 minute time limit and standardized instructions in the front of each test booklet (Lamke and Nelson, 1957). This instrument was chosen because of its correlation (.737) with the 1937 edition of the Revised Stanford-Binet Test, its relative popularity, the ease of administration and scoring, and its congruence with current school testing practices. This form of the instrument was designed to be used with third-, fourth-, fifth-, and sixth-grade children, with odd-even reliability coefficients of .961, .953, .968, and .936 respectively (Lamke and Nelson, 1957). Although the test-retest reliability of Form A was not reported, the alternate-form reliability of Form A and Form B, administered to the same
pupils with a four to ten day interval between successive testings, was .912 with Form A taken first and Form B taken second, and .936 with Form B taken first and Form A taken second (Lamke and Nelson, 1957). Therefore, successive testings of Form A on the same pupils should result in even higher reliability coefficients.

The standard error of measurement has often been considered a more meaningful indicator of reliability than the commonly reported correlation coefficient, because the magnitude of change or the width of the true score interval is shown. Assuming a pupil takes the same test numerous times, two-thirds of the time his obtained score will not differ from his true score by more than the amount given by the standard error of measurement. The standard errors of raw scores for the Henmon-Nelson, Form A, were 3.81, 4.35, 3.53, and 4.68 (Lamke and Nelson, 1957) raw score points for third-, fourth-, fifth-, and sixth-grade children respectively. By extrapolation, the standard error of IQ scores was found to be approximately five IQ points.

Design

The independent variables under consideration were the amounts of incentives promised the Ss for improved scores on the second administration of the Henmon-Nelson test, the IQ levels of the Ss, and their sex. More specifically, the treatment conditions (incentives) under consideration included a control group promised no money, a treatment group promised ten cents, a treatment group promised twenty-five cents, and a treatment group promised fifty cents for improved scores on the second administration of the Henmon-Nelson Test of Mental Ability, Form A. The IQ levels under study were high IQ (98 and above) and low IQ (97 and
below), split at the obtained mean for the sample. The males and females were also studied as a third independent variable. The change in IQ score between the first and second administration of the Henmon-Nelson test was the dependent measure under consideration.

**Procedure**

One-hundred-one third-, fourth-, fifth-, and sixth-grade students were given the Henmon-Nelson Test of Mental Ability, Form A. The test was given separately to each class in the following manner: third-grade class first, fourth-grade class second, fifth-grade class third, and sixth-grade class last. All Ss were given the standardized instructions of the Henmon-Nelson Test of Mental Ability, Form A. After all the tests were completed and scored, the 101 children were randomly assigned to one of the four experimental conditions in order to form a fairly stratified sample of the population in each group.

Eight days after the initial administration, the Ss were retested with the same form. Due to absences, nine of the original children were not available for retesting, so the remaining 92 Ss were utilized. Their names were written on the answer sheet and the treatment group was designated by the amount of money written in the upper right hand corner of the answer sheet. The names were called and the answer sheets distributed with the following instructions read by the investigator:

*Remember the test we had last week? Well, I want you to take the same one again today. This time, however, we're going to do things a little differently. Some of you have an amount of money, either ten cents, twenty-five cents, or fifty cents, printed in the upper right hand corner of your answer sheet, while some of you don't. Don't tell anyone how much money is there, because it's a secret. Those of you who have money printed on your answer sheets will be given that amount of money, ten cents, twenty-five cents, or fifty cents, if you*
can do better on this test than you did the first time. For those of you who do not have amounts of money printed on your answer sheets, I just want to see if you can do as well this time as you did the last time.

The standardized instructions concerning the test itself were once again given and the students began. The same procedure was used in the remaining three classrooms until all 92 had taken the test. The same sequence in previous classroom testing was utilized.

Those pupils in the incentive groups who improved their scores by five or more IQ points, the standard error of measurement, were given their respective incentives. In an effort to minimize the disappointment associated with not earning an incentive, the third-, and fourth-grade students who were in the control group or who were in the treatment groups but did not improve their scores by five IQ points were given a nickel each. Likewise, the fifth- and sixth-grade students who were in the control group or who were in the treatment groups but did not improve their scores by five IQ points were given a dime each.
CHAPTER V

RESULTS

The mean change scores of the four incentive groups, two IQ groups, and two sex groups were analyzed by the use of a 4x2x2 factorial analysis of variance. The results are shown in Table 1 and indicate that regardless of sex or initial IQ there was no significant difference in performance among the treatment conditions, between sexes, between IQ levels, and no significant interaction. Therefore, none of the seven null hypotheses were rejected.

TABLE 1

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>IQ Level (A)</td>
<td>1</td>
<td>5.44</td>
<td>.19</td>
</tr>
<tr>
<td>Incentive (B)</td>
<td>3</td>
<td>6.87</td>
<td>.24</td>
</tr>
<tr>
<td>Sex (C)</td>
<td>1</td>
<td>.21</td>
<td>.01</td>
</tr>
<tr>
<td>A X B</td>
<td>3</td>
<td>55.76</td>
<td>1.96</td>
</tr>
<tr>
<td>A X C</td>
<td>1</td>
<td>16.64</td>
<td>.58</td>
</tr>
<tr>
<td>B X C</td>
<td>3</td>
<td>27.11</td>
<td>.95</td>
</tr>
<tr>
<td>A X B X C</td>
<td>3</td>
<td>27.76</td>
<td>.97</td>
</tr>
</tbody>
</table>
The results of the present study were consistent with the results of the Klugman (1944) and Fast (1967) studies. Although Klugman (1944) employed primarily Middle Class Negro and white children while the present study utilized primarily Lower Class Negro children, no significant improvement in performance under monetary incentives was found. Fast (1967), using Lower Class and Upper Middle Class whites supported these findings. Therefore, different social classes and ethnic groups did not appear to perform significantly different under monetary incentive.

Since Klugman (1944) used monetary incentives ranging from five to fifteen cents, and Fast (1967) gave a penny for each correct response, it was assumed that greater magnitudes of money would have differential effects on performance. The results, however, did not confirm these expectations. The magnitude of the incentive employed did not appear to differentially affect performance.

Two important differences among these studies were the type of test employed and the immediacy of the reinforcement. Although Klugman (1944) used the Revised Stanford-Binet and administered the incentives immediately after the desired response, and Fast (1967) used the Wechsler Intelligence Scale for Children and administered the incentives
immediately after each correct response, the present study used the Henmon-Nelson Test of Mental Ability, a group intelligence test, and administered the incentives after the entire test was completed. If the present study had utilized an individual intelligence test and had administered the larger magnitudes of money immediately, the improvement in scores might have been greater.

Higgins and Archer (1968) used a group intelligence test and delayed reinforcement. A combination of money and food were used as incentives, and a significant improvement in performance of Lower Class children under the extrinsic incentives was found. A difficulty in interpretation, however, arises as to whether food, money, or a combination of the two was responsible for the increment in performance in the Higgins and Archer (1968) study. As a result of their design, it is impossible to analyze the effects of food and money as separate incentives and is difficult to compare the results of the present study with their results.

The method of determining incentives, however, can be compared. While the present study randomly selected various amounts of money to use as incentives, Higgins and Archer (1968) employed self-chosen incentives, ones the children selected as most valuable to them. Therefore, although the type of test used and the immediacy of the reinforcement did not seem to affect performance, the chosen incentives did. Thus, incentives probably should be defined by their value to the Ss, and not in terms of some experimental incentive the examiner assumes will motivate the S to greater performance.
There are several possible reasons why the present study did not detect a significant increase in performance under the monetary incentives. The presence of the control group and the experimental groups in the same room during the administration of the tests might have had an effect on the performance of the control Ss. Hearing the promise of money to the experimental groups might have been generalized to their own performance, thus destroying the effectiveness as a control for practice-effects. Likewise, although money has been shown to be effective in increasing performance in the classroom, its effectiveness in improving intelligence test performance has not been demonstrated. Therefore, since no increase beyond the standard error of measurement was evident for the groups in this study, it might be that money itself is not an effective incentive in increasing performance on intelligence tests.

From the results of the present study, money as an incentive in motivating intelligence test performance appears to be somewhat questionable, at least as a motivating factor in performance on group intelligence tests. The future use of incentives would probably be more effective if self-chosen by the individuals under study. It can no longer be assumed that a specific incentive might be motivating, but the reality of such an assumption must be tested before implementation. It is also possible that intelligence tests are so robust they can absorb the effects of extraneous variables, such as the use of incentives, and actually accurately evaluate an individual's level of intellectual functioning.
Further research along these lines would first entail the need for information concerning the motivational structures of Lower Class children. Then, the effects of these established motivational factors would warrant investigation.

A more valid study in this area would require the use of a much larger sample of the Lower Class population, possibly spanning an entire school district. Variables under study could include the use of incentives shown to be of value to Ss, black and white Lower Class groups, individual versus group intelligence tests, immediate versus delayed incentives, and the use of single versus a combination of effective incentives. The experimental groups would have to be separated from each other to eliminate possible contamination of the effects of the incentives.
REFERENCES


Hurlock, E. B. The effects of incentives upon the constancy of the I. Q. *Journal of Genetic Psychology*, 1925, 32, 422-434. (a)

Hurlock, E. B. An evaluation of certain incentives used in school work. *Journal of Educational Psychology*, 1925, 16, 145-159. (b)


Wilcutt, H. C., & Kennedy, W. A. Relation of intelligence to effectiveness of praise and reproof as reinforcers for fourth-graders. Perceptual Motor Skills, 1963, 17, 695-697.
