THE EFFECTS OF INCENTIVES ON CHILDREN'S TEST TAKING BEHAVIOR

by

Ronald Charles Kapp

A Thesis Submitted to the Faculty of the DEPARTMENT OF EDUCATIONAL PSYCHOLOGY
In Partial Fulfillment of the Requirements For the Degree of MASTER OF ARTS
In the Graduate College
THE UNIVERSITY OF ARIZONA

1976
STATEMENT BY AUTHOR

This thesis has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of the manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Ronald C. Kapp

APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

Thomas R. Kratochwill
Assistant Professor of Educational Psychology

Date: April 30, 1976
ACKNOWLEDGMENTS

I owe much to those individuals who helped me in this thesis endeavor. Their valuable time invested in this study is greatly appreciated. I want to thank Drs. Thomas R. Kratochwill, Glen Nicholson, and Shitala Mishra. Their helpful assistance was invaluable in undertaking this study. I wish to extend special heartfelt thanks to Dr. Thomas R. Kratochwill. Without his courageous efforts and timely suggestions, this study would have been difficult to complete.

Special appreciation is also extended to the staff of SS. Peter and Paul School, Tucson, Arizona.

Lastly, to my wonderful wife, Terri, whose abiding belief in my potential supplied encouragement and support throughout this endeavor, I owe so much more than words can express.
<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF ILLUSTRATIONS</td>
<td>v</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>vi</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Review of Previous Research</td>
<td>3</td>
</tr>
<tr>
<td>Summary</td>
<td>9</td>
</tr>
<tr>
<td>2. METHOD AND PROCEDURES</td>
<td>13</td>
</tr>
<tr>
<td>Subjects</td>
<td>13</td>
</tr>
<tr>
<td>Instrument</td>
<td>14</td>
</tr>
<tr>
<td>Procedures</td>
<td>14</td>
</tr>
<tr>
<td>3. RESULTS</td>
<td>17</td>
</tr>
<tr>
<td>Subject One</td>
<td>17</td>
</tr>
<tr>
<td>Subject Two</td>
<td>19</td>
</tr>
<tr>
<td>Subject Three</td>
<td>19</td>
</tr>
<tr>
<td>4. DISCUSSION</td>
<td>28</td>
</tr>
<tr>
<td>APPENDIX A: DEPENDENT VARIABLES</td>
<td>30</td>
</tr>
<tr>
<td>LIST OF REFERENCES</td>
<td>33</td>
</tr>
</tbody>
</table>
# LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $S_1$ On-Task</td>
<td>18</td>
</tr>
<tr>
<td>2. $S_1$ Playing</td>
<td>20</td>
</tr>
<tr>
<td>3. $S_1$ Out-of-Chair</td>
<td>21</td>
</tr>
<tr>
<td>4. $S_2$ On-Task</td>
<td>22</td>
</tr>
<tr>
<td>5. $S_2$ Playing</td>
<td>23</td>
</tr>
<tr>
<td>6. $S_3$ On-Task</td>
<td>25</td>
</tr>
<tr>
<td>7. $S_3$ Playing</td>
<td>26</td>
</tr>
<tr>
<td>8. $S_3$ Out-of-Chair</td>
<td>27</td>
</tr>
</tbody>
</table>
ABSTRACT

The number of children routinely administered standardized intelligence tests in the public schools is staggering. However, we still do not know the most effective way to establish conditions that maximize the individual child's performance. Clearly, some of the conditions that improve test performance are related to the relationship between the examiner and the child. The purpose of this study was to examine how various incentives influence individual test-taking behavior. A multi-element baseline design was utilized with one of three administration conditions applied to each subject. The individual administration conditions were: (a) verbal feedback for correct and incorrect responses, (b) monetary feedback for each correct response, and (c) praise for each correct response. On-Task and Off-Task behaviors were recorded at five-second intervals. Each administration was recorded on video tape to facilitate observer recording. The subjects consisted of three children who would typically represent referrals to the school psychologist. Children were chosen who were reported to have (a) difficulty focusing their attention to a task, (b) very non-verbal behavior during classroom activities, and (c) consistently obtained low scores on both standardized and regular curricular examinations. Results suggested differential responses to incentives by all three of the subjects.
CHAPTER 1

INTRODUCTION

Despite the number of children who are administered standardized tests, we still do not know the most effective way to establish conditions that maximize the individual child's performance during actual testing. Clearly some conditions that improve test performance are related to the relationship between examiner and child. Such things as the amount and frequency of feedback as well as the type of reinforcement are of demonstrated importance in testing situations (Terman and Merrill 1937, Wechsler 1949, 1974). However, most of our knowledge in this area is derived from research that has focused on what may be effective incentives for "groups" of children. That is, after a group of children receive a standardized test under some reinforcement condition (and demonstrate improved performance under such an incentive), it is assumed that the particular incentive is good for "all" children.

The procedure of generalizing that some form of encouragement improves effort and, hence, test performance, has also been characteristic of test authors who suggest that certain "standard" strategies (e.g., telling the child that he/she is
doing good, noncontingent on performance) should be routinely used during test administration (Terman and Merrill 1937).

Unfortunately, this approach offers little information on how to maximize performance in the individual child. In this regard, school psychologists are typically interested in the individual child's test score and how it reflects ability to profit from certain educational strategies. If it can be demonstrated that certain reinforcers are effective with certain children, the possibility of further individualizing the assessment strategy is offered.

Given that there is a need to establish conditions under which a given child may profit from certain incentives during testing, what techniques would offer the most useful way to investigate this question? First, each child must be individually tested under conditions that maximize his or her test performance. Second, these conditions must demonstrate a functional relationship to improvement in the child's attention to the task, elimination of certain disruptive test-taking behaviors, and, ultimately, to improved scores. In this regard, past research has focused almost exclusively on improved test scores and has ignored other parameters important to test performance.

The present investigation proposes to examine the relationship between the number of in-test behaviors and various reinforcers. Three children, who represent typical testing referrals to a school psychologist, were used as subjects. These
children exhibited certain behaviors that would likely militate against effective test performance and represent those test referrals in which standardized administration procedures would not be most effective in maintaining effort and involvement. Each child received the 1974 Wechsler Intelligence Scale for Children-Revised (WISC-R) in a multi-element baseline design where the treatment phase consisted of one reinforcer chosen for its presumed effectiveness in influencing the dependent measures of interest.

**Review of Previous Research**

Over the years psychologists have been interested in determining how intelligence test performance can be improved with certain responses from the examiner. Terman and Merrill (1937) suggested that the goal of a mental test should be to elicit an individual's best performance. Benton (1936, p. 495), going a step further, wrote, "If it were to be shown that motivational factors do have an appreciable effect upon the efficiency of performance in group intelligence tests, it would indicate that such factors must be controlled if test scores are to be considered valid indices of the intelligence of the children tested." To insure that the maximum possible performance was obtained on the Stanford-Binet, Terman (1937) also recommended the lavish use of praise such as "fine" or "splendid."

However, with the development of the 1949 Wechsler Intelligence Scale for Children, Wechsler's statements regarding
reinforcement on the part of the examiner seem to completely con­
tradict those of Terman. Wechsler (1949, p. 19) maintained that
because of the variability of different children's reactions to
commendations from adults, "In no case should the examiner indi­
cate dissatisfaction with a response as given nor build up an
expectancy of approval in the subject so that giving no comment
would be interpreted by him as disapproval."

In the manual for the 1974 WISC-R, the author suggested
that the child should be encouraged throughout the test to remain
on the task at hand. Judicious reinforcement of effort is en­
couraged but reinforcement of success is not (Wechsler 1974).
Through encouragement and reinforcement of effort, the author
assumes that the examinees are exhibiting their best performance.
However, in the past, studies have reported conflicting results
when such motivating factors as feedback and social praise are
utilized in conjunction with standard test administration
procedures.

Sweet and Ringness (1971), in an attempt to determine the
effects of feedback and monetary reinforcement on the intelli­
gence test scores of boys with differing socioeconomic back­
grounds, found no significant increase on the WISC verbal scaled
scores for middle-class whites and lower-class Negroes when
tested under feedback or monetary reinforcement conditions. They
did, however, find a significant increase in scores for lower-
class whites tested under feedback conditions.
In a similar study, Cook (1973) attempted to determine the effects of verbal and monetary feedback for correct responses on the WISC for Anglo and Spanish-American children. Verbal feedback consisted of telling the examinee whether each response was correct. Monetary feedback consisted of giving one penny for each correct response. The results showed that there was no significant difference in intelligence test scores that could be attributed to feedback conditions.

Other studies utilizing delayed feedback techniques with various test instruments have also reported differential results. Benton (1936), Fast (1967), Flook and Robinson (1972), and Ziv (1972) reported no significant increase in test scores when delayed feedback was administered. The length of time between tests and feedback varied from two days to four months. Other studies (Bridgeman 1974, Gaudet and Moon 1970), however, have reported significant increases in intelligence test scores when delayed feedback was employed. Bridgeman (1974) studied the effects of success feedback on the non-verbal subtest of the Lorge-Thorndike Intelligence Test. He found that subjects, who were given success feedback on a pre-test, score significantly higher on a post-test given two days later than did subjects who were given failure feedback on the pre-test.

Gaudet and Moon (1970) attempted to determine the effects of differential feedback on responses to intelligence tests. Subjects were given feedback two weeks after they were
administered the Otis Quick-Scoring Mental Abilities Test. These researchers found that the higher a subject scored on the first test, the greater was the score increase on a subsequent administration.

Another variable that has been investigated in terms of its effects on IQ scores is social praise. In a study by Isenberg and Bass (1974), both verbal and non-verbal praise were administered to subjects when they made correct responses to items on the WAIS. The authors reported that, although no significant differences could be attributed to the treatment conditions, results did reveal a trend towards the hypothesized increase in scores. In a similar study, Kratochwill and Brody (in press) obtained results that suggested that it is the amount of feedback contained in the verbal praise conditions that influences test performance.

Witmer, Bornstein, and Dunham (1971), using four WISC subtests (Arithmetic, Digit Span, Picture Arrangement, Block Design), found a significant increase in scores of children who were administered social praise over children who were given verbal disapproval. They found no significant increase in scores between children who were tested under verbal approval conditions and children who were tested under neutral conditions. Verbal approval or disapproval was administered after response to the first question of each subtest and between each subtest. Galdieri, Barcikowski, and Witmer (1972), using 10 WISC subtests
and the same procedure as Witmer et al. (1971), also found no significant difference between the verbal approval and no approval mode of test administration.

Studies have also examined the effects of social praise and token reinforcement upon performance on specific subtests of the WISC (Bergan, McManis, and Melchart 1971, Lyle and Johnson 1973). Bergan and his associates used both token reinforcement and immediate verbal praise after each correct block placement on the Block Design subtest. The results showed that boys obtained a significant increase in scores under token reinforcement but not under verbal praise. Girls showed significant improvement under both verbal praise and token reinforcement conditions. Lyle and Johnson (1973) examined the effects of both money and praise on the speed of writing and copying symbols on the WISC Coding subtest. Although no significant increase in performance was noted on the symbol copying task under incentive conditions, there was a significant increase in the speed of writing task when money and praise were utilized.

The effects of concrete or extrinsic reinforcers on test scores have also been studied quite extensively (Benton 1936, Edlund 1972, Fast 1967, Husted, Wallin, and Wooden 1971, Klugman 1944, Smeets and Striefel 1975). Fast (1967) investigated the effects of monetary reinforcement and verbal feedback on the WISC. Utilizing a test-retest design, 60 subjects were administered the test under standard conditions. Three months later the
subjects were divided into control, verbal feedback, and monetary reward subgroups. Verbal feedback consisted of the words "right" or "correct" and the monetary reward consisted of one penny for each correct response. The results showed no significant difference between groups.

The research of Edlund (1972) contradicts Fast's study. Using alternate forms of the Stanford-Binet, Edlund administered Form L to 11 pairs of subjects under neutral conditions. After seven weeks, Form M was administered with 11 subjects receiving the test under neutral conditions and 11 subjects receiving the test under reinforcement conditions. The results showed a significant increase in scores for subjects who were tested under reinforcement conditions but no significant increase in scores for subjects who were tested under neutral conditions. Reinforcement consisted of giving M&M candy for each correct response.

In an effort to determine which type of reinforcement contingency constitutes optimum test performance, Smeets and Striefel (1975) tested deaf children on the Raven Progressive Matrices under four treatment conditions: (1) end-of-session reinforcement, (2) noncontingent reinforcement, (3) delayed reinforcement, and (4) immediate contingent reinforcement. Using a pre-test/post-test design, 44 subjects were tested under various treatments. Reinforcement consisted of check marks that were later tallied and traded for candy or pennies. The results showed a significant increase in scores of subjects who were
given immediate contingent reinforcement but no significant increase in scores under the remaining treatment conditions.

In a unique case report, Miller (1969) studied an 11 year old girl whose WISC performance had declined from the "normal" (IQ = 101, full scale) to the "moderately retarded" (IQ = 51, full scale) classification over a two year period. However, the girl showed no corresponding behavioral deficits in non-testing situations. A subsequent administration of the WISC utilized incentive conditions consisting of a poker chip for each correct response which could be exchanged for a penny after all testing was completed. On the Digit Span subtest it was determined that the incentives were not strong enough so a subsequent Digit Span was administered with the subject receiving as many chips as were numbers in each series if she repeated the series without error. Under these conditions she increased her score on the Digit Span subtest from 0 to 12. The results of the WISC administration under incentive conditions showed an increase in the full scale IQ to 106.

Summary

Many studies have examined the effects of various types of reinforcers on intelligence test performance. Incentives such as feedback (Benton 1936; Cook 1973; Fast 1967; Flook and Robinson 1972; Kratochwill and Brody, in press; Sweet and Ringness 1971; Ziv 1972), social praise (Bergan et al. 1971; Galdieri et al. 1972; Isenberg and Bass 1974; Witmer et al. 1971), and
extrinsic rewards (Ayllon and Kelly 1972; Benton 1936; Edlund 1972; Fast 1967; Husted et al. 1971; Klugman 1944; Smeets and Striefel 1975) have been studied with differential results. These studies have employed different methods of presenting the various reinforcers. Some researchers have presented reinforcers immediately, contingent upon each correct response (Cook 1973; Edlund 1972; Fast 1967; Smeets and Striefel 1975; Sweet and Ringness 1971), whereas others have presented reinforcers after the first response on each subtest and between subtests (Galdieri et al. 1972; Witmer et al. 1971); still others have delayed the presentation of any kind of reinforcer from two days to four months (Benton 1936; Bridgeman 1974; Flook and Robinson 1972; Gaudet and Moon 1970; Ziv 1972).

Based on the results of these studies, it is difficult to suggest how certain reinforcers affect intelligence test performance. However, the strongest indication of improved test scores appears when extrinsic rewards are employed immediately upon the elicitation of each correct response (Edlund 1972).

Virtually all the investigations in the IQ reinforcement area have used a large N, between group research strategy. This has been helpful in establishing the generality of various reinforcement strategies and it has demonstrated that such reinforcers can influence test scores. The extremely variable results found in research over the past 50 years, however, suggests that there are highly unique and idiosyncratic responses to incentives that
are only subsequently identified as reinforcers. In the operant tradition, this would be an expected finding.

Unfortunately, an examination of individual response to reinforcers, with one exception (e.g., Miller 1969), has essentially been ignored in the scientific literature. This is even true in studies that were published in strict operant journals such as the Journal of Applied Behavior Analysis (e.g., Ayllon and Kelly 1972; Edlund 1972; Smeets and Striefel 1975). Although the N = 1 research strategy has a strong tradition in behavioral psychology (Browning and Stover 1971; Hersen and Barlow, in press; Sidman 1960) and its use in assessing functional relationships between treatments and behavior is of demonstrated importance (Kazdin 1973), it has not been used to examine how incentives influence test behavior on such standardized ability measures as the WISC-R.

There appear to be two somewhat conceptually related approaches to the problem. One approach is to repeatedly measure the intelligence test performance of a single subject over some period of time, while monitoring the effects of some incentive. This approach offers the possibility of examining test score change as well as behavioral change in the subject.

Another possible way to examine this issue is to administer one subject one test, during which administration conditions are varied. This approach would not allow an examination of actual test score change, but would be open to an examination of
in-test behavior patterns. It is this alternative that will be addressed in the present study.
CHAPTER 2

METHOD AND PROCEDURES

Subjects

Subjects for this study were three children in grades one and two from a school in Tucson, Arizona. Subjects were chosen by the examiner and teacher. Children were chosen who were reported to have had (a) difficulty focusing their attention to academic tasks, (b) non-verbal behavior during normal classroom activities, and (c) consistently low scores on both standardized and regular curricular examinations.

Subject one \((S_1)\) was a six year old female who was frequently off-task within the classroom environment. Subject two \((S_2)\) was an eight year old male who was reported as a very quiet, withdrawn child that would not verbally interact with his peers or the classroom teacher. Subject three \((S_3)\) was a six year old female who was reported as an under-achiever in all academic activities. No previous intelligence test data were available on any of the subjects. The academic achievement of subjects one and three was reported as below average. The academic achievement of subject two was considered average by the classroom teacher.
**Instrument**

The 1974 Wechsler Intelligence Scale for Children-Revised (WISC-R) was the instrument used for this project. This intelligence test was chosen because of an increasing trend in the use of this instrument by school psychologists as the primary method of assessing children who have been referred because of various behavior or learning problems.

The WISC-R consists of 12 subtests, of which six are designed to assess a child's performance IQ and six are designed to assess a child's verbal IQ. A full scale IQ score is also available. The test is designed for the assessment of children between the ages of 6 years, 0 months and 16 years, 11 months.

**Procedures**

In an effort to determine the effects of the various incentives on the subjects' in-test behavior, a variation of the multi-element baseline design was used (Ulman and Sulzer-Azaroff, 1975). In this design the first subtest served as a baseline condition, followed by two subtests that served as the treatment condition. The next two subtests again served as the baseline condition. This alternating method was used until all 12 of the subtests were administered. This method of research was chosen because the examiner believed it would provide the best indication of any functional relationship between the subjects' in-test behavior and the treatment. In this way there were totals of six baseline conditions and six treatment conditions for each
subject. Furthermore, three verbal subtests and three performance subtests served under both baseline and treatment phases.

One of three different treatment conditions were employed with each subject. The three treatment conditions consisted of (1) feedback for correct and incorrect responses, or (2) praise for correct responses, or (3) a monetary reward for correct responses. The treatments were further operationalized as follows:

1. **Feedback.** In this treatment the child received verbal feedback for each correct and incorrect response. The verbal feedback consisted of the words "correct" when a response was correct and "incorrect" when a response was incorrect. In addition to verbal feedback, the subject also had a pencil and piece of paper available on which to record the number of correct and incorrect responses made.

2. **Praise.** In this treatment the examiner gave verbal praise to the subject for each correct response. Verbal praise consisted of such statements as "very good," "great," or "you're really doing well!"

3. **Monetary Reward.** In this treatment the subject received a penny for each correct response. The pennies were to remain on the table in full view of the child throughout the test session.
To determine the effects of the various treatments on the subjects' behavior, a number of dependent variables were examined. These are centered around a subject's on-task and off-task behaviors while in the testing situation. A complete list of the potential dependent variables and their operational definitions can be found in Appendix A.

Each test was videotaped to facilitate scoring test related behaviors. Scoring was done by two independent observers using electronic timing devices set at five second intervals. Every other five second interval was scored and the behaviors were checked off on a coding sheet. Training of the observers was accomplished by having them view a standard WISC administration on videotape until a relatively high interobserver agreement was obtained. Interobserver agreement during the experiment was determined by calculation of Kappa (Hartmann, in press). This was chosen because of its ability to indicate the proportion of agreements, corrected for chance agreements. Kappa for the experiment was .82.
CHAPTER 3

RESULTS

The sections below represent the results of the particular treatment conditions for each subject. The effect of each treatment on the three dependent variables that occurred most frequently is discussed. Although the remaining seven dependent variables of interest were noted during the experiment, their frequency of occurrence was at best minimal, and, therefore, are not discussed. The three dependent variables of interest are: (1) On-Task, (2) Playing, and (3) Out-of-Chair. Results are discussed in terms of percent of occurrence for each behavior.

Subject One

Subject One revealed a dramatic decline of On-Task behavior from the beginning of the test to the end (Figure 1). The mean On-Task behavior under neutral conditions was 68%. The mean On-Task behavior under treatment conditions (pennies) was 42%. The highest On-Task behavior (100%) occurred under treatment conditions on the Similarities subtest. The lowest On-Task behavior (0%) also occurred under the treatment condition on the Object Assembly subtest.
PERCENT OCCURRENCE

SUBTESTS

--- pennies

--- standard

Figure 1. S₁ On-Task
Subject One exhibited very little Playing behavior throughout the entire test administration (Figure 2). The mean Playing behavior under neutral conditions was 2% and the mean Playing behavior under the treatment condition was 4%.

Subject One increased her Out-of-Chair behavior dramatically as the test progressed, regardless of what condition was being incorporated (Figure 3). The mean Out-of-Chair behavior under the neutral condition was 20% and the mean Out-of-Chair behavior under the treatment condition was 53%.

Subject Two revealed almost no variability of On-Task behavior under neutral conditions (Figure 4). The mean On-Task behavior under neutral conditions was 99%. On-Task behavior under treatment conditions (feedback) revealed much scatter between subtests. The mean On-Task behavior under treatment was 74%.

Subject Two exhibited much more Playing behavior under the treatment condition than under the neutral condition (Figure 5). The mean under the neutral condition was only 2% while the mean under the treatment condition was 26%.

Subject Two did not exhibit any Out-of-Chair behavior throughout the entire test administration.

Subject Three revealed considerable variability of On-Task behavior under both neutral and treatment conditions
Figure 2. S1 Playing
Figure 3. S₁ Out-of-Chair
Figure 4. S2 On-Task
Figure 5. S₂ Playing
(Figure 6). The mean On-Task behavior under neutral conditions was 60%. The mean On-Task behavior under treatment conditions (praise) was 42%.

Subject Three also exhibited highly variable Playing behavior throughout the entire test administration (Figure 7). The mean Playing behavior under the neutral condition was 18% and the mean Playing behavior under the treatment condition was 25%.

Subject Three again exhibited highly variable Out-of-Chair behavior (Figure 8). The mean Out-of-Chair behavior under the neutral condition was 13% and the mean Out-of-Chair behavior under the treatment condition was 33%.
Figure 6. $S_3$ On-Task
Figure 7. S₃ Playing
Figure 8. S₃ Out-of-Chair
CHAPTER 4

DISCUSSION

The results of this study suggest that the relationship between the child and examiner during an IQ test administration is complex. Specific reactions from the examiner influenced the subject's test taking behaviors, but it remains unclear as to how this might affect test scores. One might expect that with the contingent application of an incentive, an increase in on-task behavior and a corresponding decrease in off-task behaviors might result. In this study, however, an occasional increase in off-task behaviors was noted. One explanation for the paradoxical results may rest in the techniques employed by the examiner in administering the various incentives. For example, the possibility exists that the incentive was administered contingent upon a correct response but also concurrent with an off-task behavior. This could, in effect, reinforce elicitation of a correct response and off-task behavior.

The findings that these two types of behaviors can be independent has some support in the literature (cf. Ferritor et al. 1972). Ferritor et al. (1972) found that performance contingencies increased the percent of correct problems, but a concomitant decline in attending and an increase in disruptions occurred.
This study and the present findings suggest that behavioral contingencies do not always have the positive effect on performance that is often implied.

There is also the possibility that the particular treatment chosen for each student was not a sufficiently motivating incentive for the subject. Viewed in this respect, it can be concluded that the treatments had no effect on the subject's test taking behavior and that they were possibly just exhibiting their everyday classroom behavior.

It is recommended that further research in this area be undertaken. Various design strategies could be utilized in an attempt to determine whether a functional relationship between incentives and test taking behaviors exist. Also an extensive reinforcement survey of each individual subject should be carried out before actual testing takes place.
APPENDIX A

DEPENDENT VARIABLES

1. On Task (T)

   **Purpose**: To detect attending behaviors that could be considered conducive to good test performance.

   **Description**:
   
   A. Subject's visual attention is directed toward the examiner.
   
   B. Subject's visual attention is directed toward the test material.

2. Playing (P)

   **Purpose**: Playing is intended to monitor often subtle manipulative behavior that is distracting to the child.

   **Description**:
   
   Child uses his hands to play with his own or community property, so that such behavior is incompatible (or would be incompatible) with test performance.

   **Includes**: Playing with comb or wallet. Cleaning nails with pencil. Drawing on self. Manipulating pencil in such a manner as to make the behavior
incompatible with test performance, e.g., shoving pencil back and forth on desk; waving pencil through air as an airplane. Picking scabs, nails, or nose if the desired "object" is separated from the body and manipulated.

Excludes: Touching other's property. Playing with own clothes. Note: Include if article is removed from body, e.g., shoes, tie buttons, scarf, etc., and is manipulated. Lifting desk or chair with feet.
Random banging of pencil on desk. Simple twiddling of pencil, banging pencil, or putting pencil in mouth, hair, behind ears, etc., if child attends to such behavior and ceases attending to assigned task.

3. Out-of-Chair (OC)

Purpose: Out-of-Chair is intended to monitor the gross motor behavior of the child removing himself from his seat entirely. Such behavior may interfere with the child's test performance.

Description:

Observable movement of the child from his chair when not permitted or requested by examiner. None of the child's weight is to be supported by the chair, but the child may be in physical contact with the chair.

Includes: Child is leaning on desk and has either lost all contact with the chair or his weight is
actually being supported by the chair. When child is full standing and the back of legs touch chair, or child is full standing and is touching back of chair with hands. Stretching (if child actually leaves seat).

Excludes: Retrieval of an accidentally dropped test-related object. Leaning forward to pick up a test-related object even if all contact with the chair is momentarily lost, provided the child is not standing fully erect on feet.

Seven additional dependent variables were considered in this study. However, their frequency of occurrence was insignificant and, therefore, not discussed in the study. The seven additional dependent variables were: (4) Non-Compliance, (5) Physical Destruction, (6) Solicitation of Examiner, (7) Vocalization, (8) Orienting Response, (9), Noise, (10) Aggression.
LIST OF REFERENCES


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.


