MARTIN, Robert Allan, 1937–
INDIVIDUAL REWARD PREFERENCES AND EFFORT
ON SCHOOL TASKS.

The Ohio State University, Ph.D., 1971
Psychology, general

University Microfilms, A XEROC Company, Ann Arbor, Michigan
INDIVIDUAL REWARD PREFERENCES
AND EFFORT ON SCHOOL TASKS

DISSERTATION
Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Robert Allan Martin, B.S., B.M., M.A.

* * * * * *

The Ohio State University
1971

Approved by
Donald C. Smith
Adviser
Department of Psychology
ACKNOWLEDGEMENTS

The author expresses his appreciation to Dr. Peter Dunn-Rankin and Dr. Carol Cartwright for making available the results of their research with the Reward Preference Inventory. Appreciation is also expressed to Professors Donald Bersoff, John Cooper and Thomas Stephens for assistance during various phases of the current study. The author owes a special debt of gratitude to Professor Donald C. Smith for guidance and encouragement throughout ten years of graduate study. The author owes his greatest debt of gratitude to his wife, Kim, for the endless years of patient sacrifice and encouragement that made completion of this study possible.
VITA

December 22, 1937 .......... Born - Cleveland, Ohio

1960 ....................... B.S. Educ., B. Music
The Ohio State University
Columbus, Ohio

1960-61 .................... Director of Instrumental Music
Firelands Local Schools
Oberlin, Ohio

1963-64 .................... Intern School Psychologist
Lancaster City Schools
Lancaster, Ohio

1964 ....................... M.A., The Ohio State University
Columbus, Ohio

1964-66 .................... School Psychologist
Lancaster City Schools
Lancaster, Ohio

1964-66 .................... Instructor, Ohio University
Lancaster, Ohio
Lockbourne Air Force Base, Ohio

1966-71 .................... Chief School Psychologist
Lancaster City Schools
Lancaster, Ohio

PUBLICATIONS

Use of learning cues with the Bender Visual Motor Gestalt Test
in screening children for neurological impairment. *Journal
of Consulting Psychology*, 1967, 31, 205–209. (with
Donald C. Smith)

Survey of school programs for neurologically handicapped children
in Ohio. Monograph, Ohio School Psychologists Association,
1969.
FIELDS OF STUDY

Major Field: School Psychology. Professor Donald C. Smith


Studies in Exceptional Children. Professors Reginald Jones, Donald Cavin, David Sabatino
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF CHARTS</td>
<td>ix</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>I. DEVELOPMENT OF THE PROBLEM</td>
<td>1</td>
</tr>
<tr>
<td>Related Literature</td>
<td></td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td></td>
</tr>
<tr>
<td>II. PROCEDURES</td>
<td>40</td>
</tr>
<tr>
<td>III. RESULTS AND DISCUSSION</td>
<td>60</td>
</tr>
<tr>
<td>Results</td>
<td></td>
</tr>
<tr>
<td>Additional Analyses</td>
<td></td>
</tr>
<tr>
<td>Discussion</td>
<td></td>
</tr>
<tr>
<td>IV. SUMMARY AND CONCLUSIONS</td>
<td>109</td>
</tr>
<tr>
<td>Conclusions</td>
<td></td>
</tr>
<tr>
<td>APPENDIX</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>117</td>
</tr>
<tr>
<td>B</td>
<td>119</td>
</tr>
<tr>
<td>C</td>
<td>124</td>
</tr>
<tr>
<td>D</td>
<td>126</td>
</tr>
<tr>
<td>E</td>
<td>127</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>Page</td>
</tr>
<tr>
<td>------------------</td>
<td>------</td>
</tr>
<tr>
<td>F</td>
<td>128</td>
</tr>
<tr>
<td>G</td>
<td>130</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>135</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table | Page
--- | ---
1. Assignment of Subjects from Classes to Groups | 44
2. Descriptive Data for Each Group and Entire Sample | 47
3. Experimental Conditions for Each Group | 54
4. Chi-Square Contingency Table for Hypothesis I | 62
5. Means and Standard Deviations of Experimental and Control Subject Performance on Base Rate Task and Experimental Task | 63
6. Chi-Square Contingency Table for Hypothesis II Entire Sample | 67
7. Chi-Square Contingency Table for Hypothesis II Class 5 Results Eliminated | 69
8. Chi-Square Contingency Table for Hypothesis III All Subjects Eligible for Reward | 72
9. Chi-Square Contingency Table for Hypothesis III Class 5 Results Eliminated | 73
10. Summary of Separate Analyses of Hypotheses IIIB, Class 5 Data Eliminated | 75
11. Chi-Square Contingency Table of Results of Comparative Ability of Most and Least Preferred RPI Categories to Evoke Effort 50% Beyond Base Rate, Class 5 Results Eliminated | 77
12. Category Choices for Each Group and the Entire Sample | 81
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.</td>
<td>Frequency of Most and Least Preferred Reward Categories for Each Group and the Entire Sample</td>
<td>82</td>
</tr>
<tr>
<td>14.</td>
<td>Frequency of Reward Category Selection by Subjects Eligible for a Reward</td>
<td>83</td>
</tr>
<tr>
<td>15.</td>
<td>Predictability of the Reward Preference Inventory and of Each RPI Category among Experimental Groups and Classrooms</td>
<td>85</td>
</tr>
<tr>
<td>16.</td>
<td>Strength and Discrepancy Analyses for Consumable Reward and Adult Approval Categories</td>
<td>88</td>
</tr>
<tr>
<td>17.</td>
<td>Comparison of Predicted Strength and Actual Strength of Consumable Reward and Adult Approval Categories in All Classrooms</td>
<td>91</td>
</tr>
<tr>
<td>18.</td>
<td>Strength and Discrepancy Analyses for the Peer Category</td>
<td>96</td>
</tr>
<tr>
<td>19.</td>
<td>Intercorrelations among Achievement, Effort and Perseverance</td>
<td>98</td>
</tr>
<tr>
<td>20.</td>
<td>Intercorrelations among Reward Preference Inventory Categories</td>
<td>99</td>
</tr>
<tr>
<td>Chart</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>I. Flow Chart of Methodology and Procedures Used with Each Classroom</td>
<td>56</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER I
DEVELOPMENT OF THE PROBLEM

Fostering individual academic achievement is a primary goal of education. Research has shown this to be affected by many factors. Chief among these have been intelligence and effort. Some studies have indicated that intelligence accounts for only a portion of the variance in achievement among elementary school children. (Bruner, 1966; Gagne, 1967). Effort has been thought to account for much of the remainder.

Virtually every classroom has included students achieving below capability because of inadequate effort upon academic tasks. Functional analysis of behavior has demonstrated that task persistence (effort) can be affected by reinforcement. Despite systematic study of conditions that evoke and maintain effort upon academic tasks in classrooms, most teachers still rely upon "traditional" practices which frequently provide inadequate types or contingencies of reinforcement for particular individuals. As a result many students have failed to realize their academic potential.

Programmed learning and behavior modification studies have demonstrated the effects of varying reinforcement contingencies for children (Smith and Moore, 1962; Ullman and Krasner, 1965). Other research has shown that individual reward preferences differ and that these differences affect performance (Brackbill and Jack, 1958;
Heber, 1959). Several investigators have suggested or explored methods to determine effective rewards for individual children (Dunlap, 1953; King and Dunn-Rankin, 1965; Premack, 1965; Shutz and Naumoff, 1964; Stephens, 1970; Tharp and Wetzel, 1969; Witryol et. al., 1960, 1962, 1965). The methods of assessment used in these studies can be categorized as either direct choice or indirect choice.

In direct choice methods several reward objects are placed before the subject for his selection. In indirect choice methods the subject indicated a preference among rewards not physically present. Indirect choice permits a wider selection of rewards including intangibles such as adult approval, peer approval or independence. Despite potential advantages of indirect reward assessment for classroom use, few such techniques have been developed.

King and Dunn-Rankin (1965) constructed a forced-choice inventory (Reward Preference Inventory) to sample children's preferences in five categories - adult approval, competition, consumable, peer approval and independence. The Reward Preference Inventory (RPI) appears to be the only instrument available to assess children's reward preferences in classrooms that has been developed in a thorough, systematic manner. The RPI is efficient to administer individually or in groups. It has been revised several times and some normative data has been gathered (Dunn-Rankin and Shimizu, 1968). However, only one study of its utility in increasing children's
effort upon academic tasks has been reported (Cartwright, 1968). This study contained several methodological problems that made it an inadequate evaluation of the RPI for this purpose. (The methodological problems will be discussed later in this paper.) Therefore, appropriate methodology is needed to evaluate the utility of the RPI and similar inventories in controlling children's academic effort in classrooms.

Related Literature

**Achievement and Effort**

Intelligence has been recognized to be a key factor in achievement. However, its effect upon achievement has resulted from its interaction with many other factors. Gagne (1967) edited a collection of papers from a conference on individual differences and learning. In one of these House and Zeaman reviewed numerous studies suggesting that intelligence had less effect upon learning than was generally believed. They concluded further that lack of attention to relevant details in learning tasks differentiated retardates from normals rather than speed of learning. Retardates were thought to suffer from "attention deficits".

In another paper from this conference Cronbach (1967) proposed adaptation of instruction to accommodate individual differences. Differential methods for handling differing motivational types were discussed. Cronbach suggested that constructively motivated students
(success strivers) needed less structure, less feedback and less encouragement that defensively motivated students (failure avoiders).

Bruner (1966) edited a collection of papers from the "Working Conference on Children's Learning". They focused on three areas – evoking and maintaining a child's attention and effort upon learning tasks, generalization of learning from one area to another and sequencing materials for effective learning. The papers on evoking and maintaining a child's effort and attention were the most numerous. In one of these, Stevenson reviewed literature on persistence. He stated that a child persisted at a task when it brought him desirable (in his view) consequences. Stevenson felt that a child's view of desirable consequences (reinforcement values) was influenced by his social class, social deprivation level, age and sex. He felt also that effort was promoted by an appropriate task with clear sub-goals and clear consequences for persistence or non-persistence.

Heckhausen (1967) reviewed extensively the literature on achievement motivation. He reported several studies which demonstrated a relationship between task persistence and achievement motivation in children (Heckhausen and Roelofsen, 1962; Winterbottom, 1958) and adults (French and Thomas, 1958; Thomas, 1956). These lead him to conclude that "Highly motivated Ss work for a much longer time than Ss with low motivation do on a complex task in which feedback about the accuracy of the proposed solutions
is absent; they also perform better regardless of the time taken for the task" (1967, p. 113).

Later in this review Heckhausen reported numerous studies relating high achievement motivation to success in school or college (Atkinson and Litwin, 1960; Heckhausen, 1963; McClelland et al., 1953; McKeachie, 1961; McKeachie, et al., 1955; Meyer et al., 1965; Robinson, 1964; Rosen, 1956; Shaw, 1961; Uhlinger and Stephens, 1960; Weiss et al., 1959).

He stated that, "One ought also to consider the mediating dimension of reinforcement for accomplishment when he seeks the connection between motivation and achievement in school, college, or career" (1967, p. 131).

Effort and Reinforcement

The evidence relating effort to academic achievement has suggested that a teacher's primary responsibility should be to evoke and maintain children's effort upon appropriate learning tasks. Gagne (1965) shared this view.

It has frequently been said that motivating the student is one of the most important jobs the teacher can do. The present description of educational decisions leads to the view that motivation should probably be considered the primary task of the teacher. Certainly it is the most important function that the 'man on the other end of the log' can undertake. Learning itself takes place within the learner, largely as a result of his activity. But something must drive this activity and direct it toward the goal of achieving objectives. (1965, p. 247)
Programmed instruction has shown the effect of reinforcement upon academic effort. Skinner discussed the importance of reinforcement to academic effort.

Among the observable things which seem relevant are the consequences of studying or, roughly speaking, what the student "gets out of" studying. At one time we should have spoken of his reasons for studying or his purpose; but reasons and purposes are simply aspects of the field of operant conditioning, and our question really comes down to this: What reinforces the student when he studies? (1968, p. 146)

Many studies in programmed learning have dealt explicitly with different types and contingencies of reinforcement (Goldbeck and Briggs, 1962; Moore and Smith, 1961; Moore and Smith, 1962). However, reinforcement is implicit in nearly all programmed learning research. Pressey (1950) stressed the need for immediacy of reinforcement throughout his work on self instruction and automatic scoring of objective tests. Skinner built his approach on effective reinforcement contingencies (1954, 1968). Reinforcement is inherent also in studies of step size and cueing (Smith and Moore, 1962) and branching and individual differences (Coulson and Silberman, 1961; Crowder, 1959).

Behavior modification research also has demonstrated that effort can be controlled by reinforcement. Bijou (1965, pp. 72-74) reported success in using reinforcement to control the social and academic behavior of young institutionalized retardates. Hewett (1968) and Stephens (1970) have developed methods to control the social and academic behavior of learning and/or behaviorally handicapped children in special public school classrooms.
Becker, et. al., (1967) Hall, et. al., (1968) and Madsen, et. al., (1968) have demonstrated control of social and academic behavior of "normal" children in regular primary level classrooms while McAllister, et. al., (1969) has reported similar control in a regular secondary classroom. Bushell, et. al., (1968) has achieved similar results in a preschool classroom. Although these studies were designed primarily to eliminate disruptive and inappropriate classroom behaviors - speaking out, distracting others, inattention, dawdling - they succeeded in replacing these with other behaviors consistent with academic effort, i.e. attention to assigned work, orientation to task locus, proper participation in class discussion.

Definition

At this point a clear definition of reinforcement should be attempted. Deese stated:

A reinforcement is any stimulus that will increase or maintain the strength of a response or stimulus-response connection associated with it. A reinforcement in the case of operant behavior is a stimulus that follows the occurrence of the response and rewards it; a reinforcement in respondent behavior is the stimulus that elicits the unconditioned reflex. (1958, p. 16)

Kendler's definition illustrated some unsettled issues concerning reinforcement.

Some psychologists believe that the reinforcing properties of positive and negative reinforcers stem from their functioning as rewards. Others disagree, believing that the primary function of reinforcers is to provide an organism with information about its environment. Still others believe that the reinforcing quality of the unconditioned stimulus in classical conditioning stems from its ability to evoke the
unconditioned response. To complicate the picture still more, there are those psychologists who believe reinforcers operate in one way in classical conditioning but in another way in other learning situations. And there are still other psychologists who feel that a good deal more information is needed before this question can be answered. (1963, p. 157)

In this paper the definitions, synonyms and distinctions surrounding the term reinforcement were consistent with Cartwright's.

Terms such as reinforcement, reinforcer, reward, and incentive are often used synonymously, but there are subtle differences between incentive and reinforcement. In this paper, the terms reinforcement, reinforcer, reinforcing conditions, and reward will be used interchangeably and will refer to the stimulus that increases the probability of the occurrence of a response in a particular situation. The term incentive denotes a construct which represents expectancy of reinforcement (Logan, 1960; Rotter, 1954) and will be used as such in this paper. When a child is promised a certain stimulus upon successful completion of a task, the child's expectation that he will eventually receive that stimulus is his incentive. When the stimulus is finally presented, and if the probability of recurrence of the behavior increases, the presentation of the stimulus constitutes reinforcement. (1968, p. 4)

Requirements for Effective Reinforcement

Two conditions have been essential for reinforcement to be effective in increasing effort or controlling other aspects of behavior. First, the contingencies must be arranged so that the subject must emit the specific behavior before receiving reinforcement. Second, the reinforcer used must be valued by the subject in the type of situation where control is desired since a subject's reinforcement values may vary from one situation to another.

Proper Contingencies for Receipt

Programmed Learning. Programmed learning and behavior
modification were designed to control learning and social behavior through reinforcement. Pressey's (1950) work with automatic self-scoring devices used immediate rather than delayed reinforcement. Pressey's "teaching machine" informed the student of the accuracy of his response and, if wrong, guided him to the correct answer. Pressey felt that this corrected the greatest deficiency in classroom learning, i.e. delayed feedback, and thereby prevented repetition of incorrect responses. He believed also that a "teaching through testing" design could accelerate the training of great numbers of men during wars or national emergencies.

Skinner (1954) advocated immediate reinforcement along with careful arrangement of reinforcement contingencies (relationships between behavior and its consequences) to produce complex sequences of behavior through successive approximation. He referred to this process as "shaping". Competitive and cooperative behaviors were produced in pigeons through shaping. Pigeon "ping pong" games illustrated in elementary psychology texts (Kendler, 1963) exemplify the shaping of competitive behavior. Complex behavior sequences in pigeons have been maintained for long periods through intermittent reinforcement schedules (Ferster and Skinner, 1957). Skinner (1954) described how contingencies arranged through principles of immediate reinforcement, successive approximation and intermittent reinforcement could be applied to classrooms to make education more effective and efficient. These principles have fostered much of the subsequent research in programmed learning and behavior modification.
Behavior Modification. A contingency that linked teacher attention to the study behavior of several children in primary level classrooms was explored by Hall, Lund and Jackson (1968). They demonstrated that the percentage of study behavior during thirty minute observational periods could be increased markedly (generally doubled, often tripled or quadrupled) when teacher attention followed study behavior and teacher ignoring followed non-study behavior. This dramatic effect was illustrated through a sequence which included a baseline period during which percentage of study behavior was tallied during "normal" classroom behavior-attention contingencies, followed by a reinforcement period during which study behavior received teacher attention while non-study behavior was ignored. The first reinforcement period was followed by "reversal" during which original contingencies were reinstated and then by a second reinforcement period. Percentages of study behavior rose and fell predictably as the contingency arrangements were varied. Periodic observations later during the school year showed that the increased study behavior was maintained. This demonstrated the control which teacher behavior could have over behavior of specific students when classroom contingencies were carefully arranged and carried out.

Similar control over the behavior of selected students within classrooms was demonstrated by Becker, et. al., (1967) and Madsen et. al., (1968). Control of the behavior of entire classrooms was
demonstrated by Bushell et. al., (1968), McAllister et. al., (1969), and Thomas et. al., (1968).

Bushell et. al., (1968) increased the study behavior of a class of twelve preschool children through a system of tokens earned by attention to assigned tasks and exchangeable daily for a "special event", i.e. a short movie, field trip, story, extra gym period, etc. When the special event was made non-contingent upon the tokens, study behavior decreased. Reinstatement of the token-special event contingency increased study behavior to prior levels.

Thomas et. al., (1968) demonstrated that disruptive behavior could be increased and eliminated in a primary level class by systematically varying teacher comments of approval and disapproval. Appropriate behavior appeared to increase and decrease along with contingent teacher approval; however, systematic rater bias and other confounding factors precluded conclusive interpretation of results. When approval for appropriate behavior was discontinued while disapproving comments for disruption continued, disruptive behaviors increased. This was attributed to peer approval for disruption filling the vacuum when teacher approval for appropriate behavior ceased. The investigators suggested also that in some individuals critical comments might increase the behaviors they were designed to eliminate.

McAllister et. al., (1969) combined praise for desired behavior
with criticism for undesired behavior to reduce several inappropriate behaviors in a secondary classroom. Multiple baselines and a control group were used in this study rather than the reversal technique. Informal observation suggested that for at least one student critical comments appeared to increase rather than decrease the behavior upon which they were contingent. This observation along with those from the study cited above (Thomas et. al., 1968) suggested a need to explore the effects upon behavior of individual differences in reinforcement values.

Value to Recipient

The value of a reinforcement to a particular subject can facilitate or hamper learning. This was discussed by Mischel (1968).

The need to assess systematically the reinforcers that will produce and maintain behavior change seems evident but sometimes is overlooked. In one programmed instruction project, for example, the investigators initially used social approval and knowledge of results as reinforcers for their pupils (Birnbrauer et. al., 1965). They soon discovered, however, that incorrect and correct answers seemed to be one and the same to the children; often the youngsters did not even look at the answers supplied by the teaching programs. As the authors pointed out, learning academic subjects had little value to these children, and more effective reinforcers had to be discovered. (p. 253)

Skinner (1968, p. 20) held that if the reinforcements inherent in school subject matter were inadequate, external rewards would need to be employed. He listed preferred activities, competition, teacher approval and aversive stimulation as possible reinforcements within the classroom.
Stephens (1970) considered reward value important to learning also.

Rewards must be considered of value to learners if they are to be effective. Until the stimulus is viewed as reinforcing by the student it will have little effect on his responses. Youngsters who do not value school marks will be unwilling to perform if these are the only incentives that are held out as rewards. (1970, p. 51)

Reinforcement value was one of three basic constructs in Rotter's Social Learning Theory (1954). Rotter's definition of reinforcement value was the degree of preference for any external reinforcement when the possibility of receiving it was equal to that for any other. According to Social Learning Theory reinforcement value interacted with expectancy to determine the potential occurrence of a given behavior in a specific situation. Rotter distinguished between the subject's experience of a reinforcement (internal) and the description of the reinforcement by an observer (external). A given reinforcement could have a general value within a group or culture, yet be valued differently by individual members of that culture. Rotter attributed these discrepancies to differing individual reinforcement histories, i.e. a particular stimulus may have been associated with positive or negative conditions in an individual's life. Its value derived from these associations.

These differences often confounded behavior prediction based upon the group value of a particular reinforcement. They made evident the need to determine effective reinforcements for different individuals.
Determination of Effective Reinforcers for Individuals

Many different methods have been explored to determine effective reinforcements for particular individuals. Two broad approaches have been used for this purpose. One attempted to correlate individual characteristics with effective reinforcers while the other investigated individual preference for different rewards.

Individual Characteristics

Varying chronological age (CA), mental age (MA) and socio-economic status (SES) have been the individual characteristics most extensively explored in relationship to effective reinforcement.

Varying chronological age. Cartwright (1968) reviewed eighteen studies of reinforcer effects and varying chronological age. She concluded that although many of the studies were not comparable because of varying subject age, reinforcer type or learning task, no one reinforcer emerged consistently as the most effective for any of the age groups studied.

Cartwright's review included McCullers and Stevenson's (1960) finding that verbal comments such as "that's fine" and "very good" paired with the delivery of a marble were more effective on a knob-pressing task with a three to five year old group of children than with an eight to ten year old group. Jones and Liverant (1960) found the same effect in a comparison of four to six year old children with nine to eleven year old children on a lever pressing task.
However, Nickell and Travers (1964) reported conflicting results from a study of the differential effectiveness of approving statements and the delivery of a marble upon a knob-pressing task with groups of children equivalent in age to pre-kindergarten, kindergarten, third grade, sixth grade and ninth grade. They found that the marble was a more effective reinforcer for all age groups. They accounted for this unexpected result through the stimulus novelty of the marble delivery.

Varying mental age. Cartwright (1968) reviewed nine studies involving variation in mental age and effective reinforcers. Most of these involved comparison of mentally retarded with non-retarded individuals. Cartwright concurred with Heber's (1964) summary of a review related to motivation of mental retardates.

The research is consistent in suggesting that even severely retarded subjects are responsive to variations in type and magnitude of incentive .... (Heber, 1964, p. 164)

Cartwright's review included findings by Bijou (1963) and Birnbrauer and Lawler (1964) that social and academic behavior of retardates could be brought under the control of tokens and social reinforcers. However, these studies, as with most behavior modification research, failed to compare the effectiveness of the different reinforcers.

Some recent research, not reviewed by Cartwright, has compared further differential reward preferences of retardates and normals.
Zigler (1966) found that behavioral differences between familial retardates and normal children of the same mental age were due to motivational differences rather than physiological or cognitive defects. One of Zigler's five major conclusions was that environmental differences between institutionalized retardates and normals lead to different values for reinforcers ranked hierarchically.

Noonan and Barry (1967) compared institutionalized retardates, non-institutionalized retardates and normals on a performance task under tangible and social reinforcement conditions. They found that non-institutionalized retardates performed significantly longer than either of the other groups and performed faster than the institutionalized retardates. Noonan and Barry attributed these results to the greater stress upon non-institutionalized retardates in coping with a normal environment. They felt that the need of these subjects for praise, acceptance and support was greater than their need for tangible reinforcement.

Haywood (1967) found that intrinsically-motivated institutionalized retardates performed better on a hole-punching task for task incentive (promise of another task) while extrinsically-motivated retardates performed better for money rewards. Haywood (1968) found overachieving EMR's to be more intrinsically motivated than underachieving EMR's. The differences between the motivational orientation of overachievers and underachievers held throughout the normal and superior intellectual ranges; however, it was greatest with
EMR's. This suggested that extrinsic reward preferences might help to increase the performance of underachievers.

Watson, Orser and Sanders (1968) found that institutionalized severely mentally retarded children preferred candy to "amusement" rewards. Black and Thomas (1966) found that adult mongoloids performed better on a two-choice discrimination task for secondary rewards (tokens exchanged later for candy) than for primary (candy immediately) or social (verbal approval) reinforcements. Gardner and Brandl (1967) reported that mildly retarded adolescents demonstrated better intentional and incidental learning on a serial learning task for social reinforcement (verbal praise) than for tangible rewards (prizes) or no reward (instruction only). Insalaco (1967) found no differences in learning among institutionalized retardates for various combinations of verbal and consumable rewards. These conflicting results supported the views of Cartwright (1968) and Heber (1964) that no consistent mental age by reinforcer interaction appeared with retardates.

Varying socio-economic status. Ten studies reviewed by Cartwright revealed contradictory evidence about the relationship between reinforcer effectiveness and socio-economic level. Terrell (1958), Terrell, Durkin and Wiesley (1959), Terrell and Kennedy (1957), Zigler and DeLabry (1962), and Zigler and Kanzer (1962) found that intangible reinforcers ("right" for each correct response) were more effective with middle class children while tangible reinforcers
tokens exchanged for prizes) were more effective with lower class children. However, Douvan (1956), Fang and Davis (1967), Kulberg (1967), Miller (1963) and Rosenhan and Greenwald (1965) failed to find significant interactions between reinforcer and social class.

Shores (1969) found that middle class underachievers performed significantly better for tangible (not specified) than for intangible (response confirmation) rewards while their lower class counterparts showed the opposite effect. Normal achievers from both classes performed slightly better for intangible than for tangible rewards.

Research which has attempted to relate reinforcer effectiveness to individual characteristics has failed to produce consistent, useful results. Therefore, other means of assessing reinforcer effectiveness are needed.

Individual Preferences

The effect that reinforcement value could have upon behavior was discussed above. Cartwright (1968) reviewed several studies that demonstrated the effects of individual differences in reward preference upon performance (Brackbill and Jack, 1958; Finley and Staats, 1967; Heber, 1959; Kintz and Pappas, 1965; and Mischel, 1961). Results of these studies suggested that individual reward preference affected performance variability among subjects.

Brackbill and Jack (1958) found that when five-year-old boys were given a choice of marbles, candy, or trinkets, the three possible rewards were chosen an equal number of times. Since both age and sex of subjects were held constant, the results suggest that preference
rather than age or sex was responsible for the choice of rewards. In addition, the finding that each of the three alternative reinforcers were chosen and equal number of times suggests that there is considerable variance in individual preference for a reinforcer. When performance of this group which was provided with a choice of rewards was compared with that of another group receiving only a candy reward, only slight differences between mean trials to criterion scores were noted; variance between the two groups was large, however. The group receiving a preferred reward showed significantly less variability. The fact that variance between groups differed greatly implies that individual preference for rewards may be obscured when groups are compared in terms of mean scores.

Data provided by Kintz and Pappas (1965) support the results given in the preceding investigation. These data indicated that subjects' preferences for different verbal stimuli were directly related to task performance. Another study which supports the notion that individual preferences should be taken into account in the selection of reinforcers was recently completed by Finley and Staats (1967). These investigators asked subjects to rate words as having either positive, neutral, or negative meaning. The results of two experiments indicated that positive evaluative meaning words strengthened a motor response, negative evaluative meaning words decreased the strength of the motor response, and the neutral evaluative meaning words had little effect on the strength of the response...

Heber (1959) assigned mentally retarded individuals to reinforcing conditions of high-incentive (highly preferred reward) and low-incentive (lowly preferred reward) and observed their performance on a simple motor task. He reported that the high incentive group performed the task better than the low-incentive group and noted a decrease in performance when the high-incentive group was shifted to low-incentive condition. Little consistency between subjects in preference ranking of the eighteen available alternative rewards was also noted:

The low inter-individual consistency which was obtained in the preference ranking of reward objects raises a question as to the desirability of arbitrary selection of a uniform reward by the experimenter in studies of this type with
mentally retarded subjects. (Heber, 1959, p. 671)

Blount (1966) studied the comparative preferences of institutionalized adult retardates for M & Ms, money, cigarettes and personal possessions. The paired comparison choices revealed no significant correlation with MA, CA, or length of institutionalization. However, larger amounts of each of the rewards were preferred to smaller amounts. The effects of low-preferences vs. high-preference rewards were evaluated with a finger maze task. Subjects receiving high-preference rewards performed significantly better (.01 level).

Siegel, Forman and Williams (1966) found that the preferences of adult retardates for consumable or object rewards consistently influenced their selection of stimuli on a two-choice discrimination task, i.e. the stimulus associated with high-preference rewards was consistently selected. Siegel, Williams and Forman (1967) found that the performance of adult institutionalized retardates on a marble-dropping task was increased when switched from low-incentive to high-incentive. Performance remained the same when subjects were switched from high to low incentive and extinguished when reinforcement was discontinued. Control groups continued on the same reward throughout (either low or high) showed steady improvement.

Siegel, Forman and Williams (1967) developed a methodology for rank-ordering incentives with adult institutionalized retardates.
They found significant consistency for 10 of 25 subjects on a first-choice measure and for 14 of the subjects on a work output measure. A two-choice discrimination task produced consistent high incentive choice in 13 of the 25 subjects.

Results of studies relating individual reward preference to performance have suggested that this is a fruitful approach to effective behavior control.

Assessment Procedures. Little research has been done on the assessment of individual reward preferences. What does exist can be clarified by division into two categories - direct choice methods and indirect choice methods. Direct choice methods involve selection of a reward that is physically present to be seen and/or handled by the subject. Indirect choice involves selection of a reward not physically present. Several techniques have been used in assessing individual reward preferences. Some have been represented in both of the above categories, others in only one. Paired comparison (forced choice) and ranking methods have been used in both direct choice and indirect choice assessment. Semantic differential scales and interviews have been used in indirect choice assessment; observational techniques have been used in direct choice assessment.

Direct choice methods were investigated by Lotsof (1951) through preferences of college students in a study that compared choices in hypothetical and real situations. His subjects chose which of a
pair of liquids they would prefer to drink in a hypothetical situation (preference stated only) and in a "real" situation (chosen liquid consumed). He found similarity of preference and response latency in both situations.

A series of studies by Witryol et. al., (1960, 1962, 1965) indicated that rewards selected by children in a direct choice, paired-comparison situation were effective reinforcers. Shutz and Naumoff (1964) however, used a paired-comparison method with kindergarten children and found conflicting results. These inconsistent results suggest a need for further exploration of direct-choice, paired-comparison methods.

Dunlap (1953) used direct choice ranking to determine toy preferences of third, fourth and fifth grade boys. Dunlap's study was concerned with changing weakly differentiated, moderate preferences rather than validating strong or weak preferences. Dunlap's findings along with those of Siegel, Forman and Williams (1967) suggested that more exploration of the utility of ranking methods was needed.

Observational techniques are another method of direct choice reinforcer assessment. These have been explored and developed by Premack (1965), used by Ayllon and Azrin (1965) with adult psychotics and described by Stephens (1970) for use in schools. Premack (1965) observed the behavior of nursery school children in naturalistic settings with a wide variety of potential reinforcements available.
He attributed high reinforcement value to objects or activities with which the child occupied himself most frequently or for the greatest amount of time. These high probability behaviors (HPB) were then made contingent upon the child's engaging in low probability (infrequent) behaviors (LPB). Premack found that frequency of low probability behavior could be increased when it was followed by allowing the child to engage in a high probability behavior.

Ayllon and Azrin (1965) used Premack's observational techniques with chronic schizophrenics to determine effective reinforcers to be made available for exchange for tokens earned through rehabilitative functions. Stephens (1970) suggested casual and structured observational techniques that could be used by teachers to determine children's reinforcement values.

Several cautions should be needed in using observational techniques for reinforcer assessment. The situation observed should be similar to the one in which control is desired; it should contain all or nearly all of the potential reinforcers to be compared; all of the potential reinforcers should be equally accessible to the subject in the situation observed. Rotter (1954) emphasized the necessity to control expectancy when measuring reinforcement value. If this is not done, measurement of reinforcement preference could be contaminated by the subject's expectation that he would be more likely to receive one reinforcement than another.
Indirect choice methods of assessing reward preference were developed because of the possibility of presenting more alternative reinforcers, particularly intangibles. These methods have included structured interviews, semantic differential scales, rating scales and forced choice inventories.

Stephens (1970) described structured interview procedures for use by teachers with individuals and groups of children. In individual interviews the child was asked to choose between pairs of potential reinforcers as rewards for academic effort. The group interview was an open-ended discussion of preferences for school tasks and reinforcers. These techniques, while flexible, were unsystematic. They relied upon teacher judgment to include all available reinforcers and to compare them systematically so that preferences could be determined. It is unlikely that a teacher could do this in student interviews without more detailed, specific directions.

Tharp and Wetzel (1969) developed a structured interview to determine effective mediators of reinforcement for individual children. Since reinforcer efficacy varied from one situation to another, mediator value was explored as a potential source of variability. The structured interview was used to enhance reward value through identification of effective reward mediators for particular children in school and home situations.

Semantic differential scales were used by Finley and Staats (1967) to evaluate the positive, negative or neutral meaning of words
to individual subjects. Words with positive meaning were found to increase frequency of a motor response while words with negative meaning decreased it and neutral words had little effect upon it. This technique would be useful for evaluating verbal reinforcement, but would not provide comparison with other types of reinforcers available in classrooms.

Rating scales, while not used directly in assessing individual reward preference, have been used in conjunction with forced choice inventories (Stephens, 1970) and in the determination of rewards to be used in forced choice inventories (Dunn-Rankin and Shimizu, 1968). With systematic development rating scales could become a useful assessment technique.

Stephens (1970) described an assessment technique which combined a rating scale with a forced choice preference schedule. The rating scale included six categories, ranging from 1= I like best, to 6= I strongly dislike. The preference schedule contained different school activities. Results of this technique would yield student preferences for school activities. Stephens description was suggestive of how an assessment technique could be developed rather than a presentation of a systematically constructed, carefully validated instrument.

King and Dunn-Rankin (1965) devised a forced choice inventory to determine the relative preferences of individual children over five categories of reinforcement available in classrooms - adult
approval, competition, peer approval, independence, consumable. Competition and independence rewards had not been explored as thoroughly as the other categories of rewards. No comparison of preference across all five categories had been reported. Dunn-Rankin and Shimizu (1968) chronicled the development of the inventory (Appendix A).

The third revision of the Reward Preference Inventory included four subtests, each containing 10 independent pairings of the five reward categories (Appendix B). The list of reward statements by category is included in Appendix C.

The short term reliability of the third revision of the Reward Preference Inventory was determined through two administrations, one week apart, to eighty-seven intermediate grade children at the University of Hawaii Elementary School. The test-retest Pearson correlation for the Reward Preference Inventory was .84. Correlations for the first, second, third and fourth subtests respectively were .65, .75, .75, .75. No information was reported about reliability of category ranks for individual students.

A study of the long term reliability of RPI category ranks was done by Clum (1970) with students in intermediate level classes (chronological age 10-14) for the educable mentally retarded. She assessed the stability of reward preference category rankings over a four month period. She found that 35% of the preference rankings were stable for subjects in the 10-12 CA range while 75% were stable
for subjects in the 12-14 CA range. Chronological age was the only individual characteristic (among sex, IQ, reading achievement and arithmetic achievement) found to relate to preference stability.

The Reward Preference Inventory is unique among methods of reinforcer assessment since its development was systematic. The RPI is systematic also in that each reward category is compared with every other category in each subtest. This provides four replications for each comparison in the current revision. The procedures for administration are structured to avoid contamination of reinforcement value by expectancy. The danger of that contamination was cited by Rotter (1954).

It should be re-emphasized here that in any instance where we attempt to measure reinforcement value, we must make inferences from some behavior of the subject. If we do not control expectancy, our measure will be of the behavior potential of the response rather than of the reinforcement value with which we are concerned. (Rotter, 1954, p. 149)

The Reward Preference Inventory is efficient because of clear, simple directions and short administration time (about twenty minutes). It can be administered to groups of older children with adequate reading skills. Scoring procedures are brief and results easily understood.

The Reward Preference Inventory is more comprehensive than most methods of reinforcer assessment since it includes five different categories of potential rewards. However, several types of reinforcement were omitted that have been found to affect children's
behavior in schools. These included individual adult attention and immediate knowledge of results. Preferences for reward mediators were omitted from the RPI, also. Tharp and Wetzel (1969) found mediator preference to influence reward efficacy. Had the Reward Preference Inventory included assessment of preferences for individual adult attention, immediate knowledge of results and reward mediators, its potential utility in schools might have been enhanced.

Evaluation is needed to determine whether the reward preferences revealed by the RPI can affect children's classroom behavior. Only one study has been reported of the Reward Preference Inventory's utility in controlling children's behavior. Cartwright (1968) studied the effects of individual reinforcement preferences of elementary school children, as measured by the Reward Preference Inventory, upon a learning task and persistence task. The hypotheses of this study were:

**Hypothesis A.** Children working under conditions of high incentive will perform significantly better than children working under conditions of low incentive on a paired-associates learning task. Performance on the paired-associates learning task is measured by the number of trials to reach criterion and number of errors.

**Hypothesis B.** Children working under conditions of high incentive will work significantly better than children working under conditions of low incentive on a perseverance task. Performance on the perseverance task is measured by the number of correct responses divided by total number of responses possible. (Cartwright, 1968, p. 25)

No significant effects upon task performance were found; however, there were several shortcomings in this study which made it inadequate
as a test of the effect of reinforcement preferences upon academic behavior. These shortcomings included:

1. Directions that did not relate reinforcement preferences to increased performance on the experimental tasks.

2. Rewards that did not correspond with those selected from the reinforcement inventory.

3. Experimental tasks that had little resemblance to school assignments.

4. Evidence that the subjects were too old and/or sophisticated for the rewards included in the inventory.

5. No base rate performance data on the experimental tasks without the rewards from the Dunn-Rankin reinforcement inventory.

The most serious of the shortcomings were the directions given by Cartwright to her subjects. She did not set the reinforcement contingencies so that receipt of reinforcement depended upon quantity of effort exerted on the experimental task. The reinforcement contingencies were set as follows:

Each subject in the high incentive group was told he would receive _____ (his most highly preferred reward) upon completion of the task. Similarly, each child in the low incentive group was told he would receive _____ (his least preferred reward) when he finished the task. (Cartwright, 1968, p. 42)

Cartwright's directions for the learning task were as follows:

Now look up at the screen (pointing). On the screen you will see more pictures like those just learned only these will be different pictures and there will be more of them. You have to
do the same thing for these new pictures, tell me the name of the one that is missing. The first time through I will name both pictures. Keep working until I tell you to stop. Are there any questions? (Cartwright, 1968, p. 39)

Directions for the perseverance task were as follows:

This is a study to see how well you can find numerals. There are many different numerals on each of these pages (pointing to set of pages). You are to look carefully at each numeral. Each time you find a numeral five (show card with numeral five printed on it) draw a circle around the numeral five like this (show card with circle drawn around a numeral five). When you finish one page, go right on to the other pages. Work as long as you can and be careful so you do not make mistakes. When you get tired of doing this and want to stop, tell me you are finished. Are there any questions?

Remember to tell me when you get tired and want to stop. (Cartwright, 1968, p. 42)

Directions and reinforcement contingencies for the learning and perseverance tasks asked the child to tell the experimenter when he wanted to stop or told the child when to stop. It was neither stated nor implied that increased effort would affect the reward received. Because of this it appeared that the subjects had no incentive to perform the experimental tasks beyond the point of minimal fatigue or task boredom. The dependent variables, however, were the quality and the quantity of performance on the experimental tasks. Cartwright's methods of analysis of the dependent variables illustrate this.

Data were collected for each subject by counting the number of trials required to reach criterion performance of two continuous and errorless trials, and by counting the number of errors recorded on the response sheet. A maximum number of sixteen trials was allowed, but no subject needed all possible trials to reach criterion performance. (Cartwright, 1968, p. 40)
Data were collected by recording the total amount of time each subject worked and by deriving a standard score for fives. The standard score was computed by dividing the total number of fives circled by the constant 600 and multiplying by 1000 to eliminate the decimal. The constant 600 was used because it represented the nearest hundred above the total number of fives circled by any child in both groups (high-incentive, low-incentive). (Cartwright, 1968, p. 43)

Separate t tests for noncorrelated samples of equal size were applied to determine the significance of the difference between the following pairs of means:

1. trials to criterion on the paired-associates tasks for high-incentive and low-incentive groups;

2. number of errors on the paired-associates tasks for high-incentive and low-incentive groups;

3. number of minutes on the perseverance tasks for high-incentive and low-incentive groups;

4. standard score for five's on the perseverance tasks for high-incentive and low-incentive groups. (Cartwright, 1968, pp. 43-44)

Comparison of the directions and the experimental conditions in juxtaposition with the methods of analysis of the dependent variables suggested that the methodology of the study did not provide an adequate test of its proposed hypotheses.

2. Rewards that did not correspond with those selected from the reinforcement inventory.

Another shortcoming in Cartwright's study was that the rewards made available to the subjects were, in several cases, determined by the ease of making them available, rather than by their correspondence to those on the Reward Preference Inventory. This discrepancy could have lessened the effect of the particular reinforcement category in
influencing subject behavior. The following excerpt illustrates this:

Sample incentives from each of the five classes were arbitrarily selected by this investigator. Efficiency of administration to individuals was the major criterion for selection of the incentives. During the learning task, then, .......... children who preferred adult approval most highly and .......... children who preferred adult approval least were told "If you do a good job, when you are finished you will get a certificate of merit showing you did excellent work." (Cartwright, 1968, p. 35)

The following is a list of items included in the inventory for the adult approval category of rewards:

Adult Approval: A grade of "A" on your paper.
A grade of "100" on your paper.
Teacher writes "excellent" on your paper.

(Cartwright, 1968, p. 73)

Another excerpt indicates this discrepancy in the peer approval category.

"If you do a good job, when you get finished I will tell your best friend how well you worked; he (she) will be proud of you," for Peer Approval preferences; (Cartwright, 1968, p..35)

Below is the list of peer approval items from the Dunn-Rankin inventory.

Peer Approval: Smartest student in class says you did better than he.
Friends ask you to sit with them.
Students ask you to be on their team.

(Cartwright, 1968, p. 73)

A similar discrepancy occurred between the independence reward made available and the independence choices on the Reward Preference Inventory. While not as serious as the direction and reinforcement
contingency deficiency, these discrepancies negated the effects of Dunn-Rankin's scaling procedures (described above) which resulted in the selection of items for each category with high value and low discrimination indices. The discrepancies could have hampered the effect of the reinforcement preferences upon the experimental task behavior of the subjects.

3. Experimental tasks that had little resemblance to school assignments.

The experimental tasks employed by Cartwright had little relevance to the normal classroom situation experienced by most children. The excerpts below describe the learning task.

Ten noun pairs previously used by Kulberg (1967) constituted the stimulus-response pairs for the paired-associates task employed in this research. The noun pairs were originally selected by Kulberg from nouns included in the Peabody Picture Vocabulary test (PPVT) ............... (p. 36)

Slides were produced by photographing pictorial representations of the nouns in the PPVT. Four complete sets of slides were prepared. Each complete set of slides consisted of ten stimulus-response pair slides and ten corresponding stimulus-only slides. Each of the four sets of slides was then programmed into a slide tray for Kodak Carousel Slide Projector .......... (Cartwright, 1968, p. 36)

The following excerpt describes the perseverance task.

The perseverance task was constructed by mimeographing pages of numerals zero through nine drawn from a table of random numbers. Copies of six different pages of numerals were prepared in this manner. The pages were then stapled together with twenty-one pages per set. (Cartwright, 1968, p. 42)
Each subject was then given the task of circling fives on these pages of random numbers until he became tired and wanted to stop. Data collected were the total amount of time the subject worked and a standard score computed from the number of fives the subject circled.

From the description of these tasks it can be seen that they bore little resemblance to normal classroom activities of elementary school children. Research in personality assessment (Mischel, 1968; Murstein, 1965; Rotter, 1954) has demonstrated the situational effects upon validity and the danger of overgeneralization. Therefore, generalization of results from a study using "novel" tasks (as compared to normal classroom activity) to elementary school situations would be unjustified.

4. **Evidence that the subjects were too old and/or sophisticated for the rewards included in the inventory.**

The age and maturity level of the subjects in this study appeared too great for some of the rewards on the Dunn-Rankin inventory to be effective. The rationale for selection of intermediate age subjects was given by Cartwright (1968, p. 27).

Subjects were selected from among intermediate grade pupils rather than primary grade pupils for several reasons. The chronological age range represented in the intermediate grades is approximately nine years through thirteen years. It is well known that individual differences increase with each succeeding year. In addition, the suggestion was advanced that children's motivational systems were fairly well set by the age of 10 ...............
The rewards included on the Dunn-Rankin inventory were as follows:

**Adult Approval:** A grade of "A" on your paper.
A grade of "100" on your paper.
Teacher writes "excellent" on your paper.

**Competition:**
Teacher tells the class your work was the best.
Teacher writes your name on board because your work was best.
Be the only one in class who could answer a question.

**Consumable:**
A soft drink.
A nickel.
A scoop of ice cream.

**Peer Approval:**
Smartest student in class says you did better than he.
Friends ask you to sit with them.
Students ask you to be on their team.

**Independence:**
Be free to play outside.
Be free to draw pictures.
Be free to look at different books.
(Cartwright, 1968, p. 73)

The inappropriateness of the rewards for the subjects' maturity level was suggested by the spontaneous refusal of the reward offered upon completion of the experimental task by seven members of the low incentive group even though the option to refuse the reward had not been offered. Cartwright did not analyze the age level of the subjects spontaneously refusing the reward, nor did she list the rewards that were refused; however, the spontaneous refusals seem to indicate the inappropriateness of the rewards for the subjects' maturity level. If some rewards were spontaneously refused, it could be assumed that other rewards might have been viewed by the subjects as not too desirable. No data were presented in this study concerning the subjects' feelings about the overall desirability of the rewards.
offered. It is possible that none of the rewards on the Reward Preference Inventory were strong enough incentives to affect the subjects' behavior.

5. **No base rate performance data on the experimental tasks without the rewards from the Reward Preference Inventory.**

Another shortcoming of this study was lack of base rate data concerning how much work might have been done on the experimental tasks by the subjects had no reward been offered. Without this information it was difficult to determine whether the reinforcements increased the subjects' efforts beyond the point of minimal fatigue or task boredom (base rate). Also, the reinforcement contingencies and the directions (mentioned above) did not seem to imply that the subject work beyond this point. Therefore, it would appear necessary to gather base rate data on each subject in a setting devoid of the rewards included on the Reward Preference Inventory to determine the effect of these rewards upon experimental task performance.

**Statement of the Problem**

Effort has been found to be a key factor in academic achievement. Individual differences in reward preference have been found to affect the amount of effort expended upon a performance task. The Reward Preference Inventory has been the only comprehensive, systematic and efficient instrument reported to assess individual differences in reward preference among children in school classrooms. Only one study has been reported of the utility of the Reward Preference Inventory in
increasing the amount of effort school children expend upon academic tasks. This study contained several methodological deficiencies that made it inadequate as an evaluation of the Reward Preference Inventory.

Behavior modification has transformed psychology and education through systematic application of reinforcement to behavior. However, behavior modification approaches have included few systematic, efficient methods to assess reward preferences. Hence the reinforcements used in behavior modification studies often have been applied indiscriminately with little regard for individual preferences. Assessment of individual reward preferences might make behavior modification approaches more efficient. Reward preference inventories are easy to construct and convenient to administer. They appear likely to become a popular method to assess reinforcement preferences. As more behavioral investigators become aware of these issues, reward preference inventories seem destined to proliferate. Therefore, methodologies to evaluate the effectiveness of reward preference inventories for various purposes and in different situations need to be developed before proliferation occurs.

This study had two main purposes: (a) to assess the utility of reward preferences, as measured by the Dunn-Rankin Reward Preference Inventory, for controlling school children's academic efforts (b) to develop methodology to evaluate reward preference inventories for similar purposes and situations. The following questions were
1. Will reward preferences, indicated by Dunn-Rankin's Reward Preference Inventory, cause school children to work on an academic task beyond their point of minimal fatigue or boredom (base rate)?

2. Will an elementary school child's most preferred reward, as determined by the Dunn-Rankin inventory, be chosen consistently over his least preferred reward when a choice of either reward is offered upon completion of an appropriate academic task that requires effort beyond his base rate?

3. Will a most preferred reward control an elementary school child's effort on an appropriate academic task beyond his base rate better than a least preferred reward regardless of whether receipt of his most preferred reward requires ...

A. more effort or ...

B. less effort than receipt of his least preferred reward?

Rationale

Question 1 needed to be answered to determine whether the rewards included in the Reward Preference Inventory could control the academic efforts of elementary school children to a significantly greater extent than minimal fatigue or task boredom. A control group and base rates were included in the methodology used to answer the question. The inclusion of the control group tested the validity of the base rate technique often used in this type of research. If verified, the base rate would provide a more efficient methodology for future research on this topic.

Question 2 was included to provide a task validation of the Dunn-Rankin Reward Preference Inventory. A negative answer to question
2 would not be sufficient reason to abandon the reinforcement inventory concept although it would suggest that revision of the Reward Preference Inventory were needed.

Question 3A was included to determine whether a child's most preferred reward could overcome his disinclination to exert the effort required to attain that reward when its attainment required more effort than the attainment of his least preferred reward. Question 3B was included to determine whether a child's most preferred reward could overcome any "need" he might have had to strive for the most difficult goal available when attainment of his most preferred reward required less effort than attainment of his least preferred reward. This question (3B) compared the relative strength of stated reward preferences with the need to strive for the most difficult goal on an academic task. It was important because Cartwright (1968) found that "high achievers" performed significantly more work on experimental tasks than "low achievers" regardless of reward value. Therefore, the inclusion of 3B was necessary to determine what proportion of subjects would work for the most difficult goal regardless of the value of the reward received. Without 3B, data on 3A would be equivocal, i.e. it would be unclear whether the subjects were striving for the reward or because of some compulsion to attain the most difficult goal regardless of reward value.
CHAPTER II
PROCEDURES

Methodology and procedures for the study were devised to answer the questions posed above. Socio-economic factors of the community and school locus of the study, rationale for selection of subjects and assignment to experimental groups, descriptive information about the subjects, data collection instruments and procedures, and data analysis procedures will be discussed in this chapter.

Locus of the Study

The study was conducted during the 1969-70 school year in public school classes for educable mentally retarded children in Lancaster, Ohio. Lancaster, a community of 35,000, is located 30 miles southeast of Columbus, Ohio, a large metropolitan complex. It is a self-sufficient community with its own industries, shopping centers and communications media. Few residents commute to Columbus for employment. In addition to economic self-sufficiency, Lancaster has a "psychological" self-sufficiency based upon its early settlement and several historically significant individuals born there.

The socio-economic pattern of Lancaster reflects bimodality characteristic of its local industry, largely glassware manufacture and its satellite activities. Several companies based in Lancaster
have numerous white collar and blue collar employees. The white collar and blue collar groups tend to live in different parts of the community and to have minimal contact. The ethnic composition of Lancaster is largely white, Anglo-Saxon Protestant. Most community residents, affluent as well as poor, have lived in Lancaster all their lives. Lancaster's few transients are among the white collar group.

The Lancaster City Schools include 8600 students in grades kindergarten through 12. One of several special programs for handicapped children includes classes for the educable mentally retarded (Intelligence Quotient 50-80). The Ohio Department of Education supports these classes and prescribes standards for their composition and operation. The twelve classes, including 240 students, culminate in work experience during the last two years of high school. Socio-economic composition of the educable mentally retarded classes is almost entirely working class and lower class white.

Selection of Subjects

Rationale. The sample for the study consisted of 80 intermediate age students from classes for educable mentally retarded children in the Lancaster City Schools. Although Cartwright (1968)

1 Ninety students constituted the total population of intermediate EMR classes in the Lancaster City Schools. Those rejected were unable to score within an acceptable range after attempts on several different base rate tasks. Base rate tasks and methods of selection will be described below.
used students from regular classes in her study several considerations lead to the use of students from educable mentally retarded classes in the current study.

1. The smaller range of intellectual functioning in students from educable mentally retarded classes in comparison with students from regular classes reduced one source of variability in accounting for experimental results.

2. Volunteer bias was avoided since all students in the educable mentally retarded classes of the Lancaster City Schools were available for the sample if needed.

3. Complete and recent individual intelligence and achievement data were available for each educable mentally retarded student along with carefully verified information concerning socio-economic status, medical history and educational history.

4. The novelty of the information gathering and experimental procedures would have less effect upon students in educable mentally retarded classes than regular classes since all students in educable mentally retarded classes have experienced frequent individual testing and some have experienced systematic reinforcement contingencies for academic effort.

5. Teachers of educable mentally retarded classes were aided by having reward preferences for every student in their classes since all students in the educable mentally retarded classes of the Lancaster
City Schools were available for use in the study.

6. More dissemination and use of the results of the study were likely among teachers of educable mentally retarded classes than regular classes since a better information dissemination network is available for teachers of educable mentally retarded classes through Educational Resource Information Centers, Instructional Resource and Materials Centers, the Council for Exceptional Children and the Ohio Division of Special Education.

7. The smaller size and greater homogeneity of educable mentally retarded classes made implementation of research findings more feasible.

8. Potential generalizability of results to students in regular classes was suggested since Cartwright and Cartwright (1969) found that older children in educable mentally retarded classes had preference profiles similar to younger normal children.

Assignment to Groups

Four groups with 20 subjects in each were needed to test the hypotheses of the study. Assignment was done by a stratified random procedure to produce an equivalent chronological age distribution in each group. This was achieved by assuring that the subjects used from each of the educable mentally retarded classes were divided proportionately among the four groups. Assignment of subjects from each educable mentally retarded class to the experimental groups is illustrated in Table 1.
Descriptive Information on Subjects. Descriptive information on each subject in the sample was gathered in the following areas - chronological age, Mental Age, Intelligence Quotient, reading achievement, arithmetic achievement, socio-economic status of parent, parent education, and effort typically expended upon arithmetic tasks in the classroom.

The chronological age and Mental Age of each subject were computed as of January 1, 1970, since this was the mid-point in time for data gathering. The Intelligence Quotient for each child was from either the Stanford-Binet or the Wechsler Intelligence Scale for Children. All of the Intelligence Quotients except two had been derived from tests given since January 1, 1968. The two exceptions were from tests given in the fall of 1967. The reading and arithmetic portions of the Wide Range Achievement Test were given to each subject in April of 1970. Grade levels and standard scores were computed for each subject in both areas.

<table>
<thead>
<tr>
<th>Classes</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>17</td>
</tr>
<tr>
<td>II</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>19</td>
</tr>
<tr>
<td>III</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>IV</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>V</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>80</td>
</tr>
</tbody>
</table>
Socio-economic status ratings for each subject were based on parent occupation and derived from the Socioeconomic Index (Duncan, 1961). The Socioeconomic Index was designed to measure socio-economic status from occupation alone so that birth or death certificates, which list occupation, but not income or education, could be used to determine socio-economic status. Education and income combined were used to construct the Socioeconomic Index since this combination had been shown to correlate highly (r = .91) with occupational socio-economic status. However, neither education nor income alone correlated highly with the Socioeconomic Index. The correlation between education and the Socioeconomic Index for males in the civilian labor force in 1950 was .567. Only 32% of the variance between occupational levels in socio-economic status was included by the correlation of .567.

Because of the low correlation between education and the Socioeconomic Index, years of parent education for each subject were recorded as another measure of socio-economic level to provide comparison with Socioeconomic Index status ratings. The correlation between Socioeconomic Index ratings and years of parent education with the current sample was .59. Another reason for recording years of parent education was the extreme variability of Socioeconomic Index ratings exemplified by a greater standard deviation (14.83) than mean (14.50) Socioeconomic Index rating for the sample.

The amount of effort each child typically expended upon
arithmetic tasks in comparison with his classmates was assessed through teacher rating. Every teacher was asked to indicate in which 20 percent of the class each student usually performed in effort expended. The directions (see Appendix E) specified that the rating was to be for effort expended rather than for actual achievement in arithmetic. Effort ratings were gathered to compare teacher perceptions of classroom perseverance with base rate and experimental task perseverance.

Table 2 indicates the mean and standard deviation of each group of subjects and the entire sample used in the study for the characteristics of chronological age, mental age, Intelligence Quotient, reading grade level and standard score, arithmetic grade level and standard score, socioeconomic status, parent education, and effort typically expended upon arithmetic.

Data Collection Instruments and Procedures

The Reward Preference Inventory, base rate tasks and experimental tasks were given to the subjects by class in accordance with the time intervals mentioned below. The data were gathered from November 1969 through February 1970.

Dunn Rankin Reward Preference Inventory. The Reward Preference Inventory (described earlier) was given to each child in the classes from which the sample of experimental subjects was selected. It was administered individually according to the directions of Dunn-Rankin (Appendix B). The child was read each paired comparison item on the
TABLE 2

Descriptive Data for Each Group and Entire Sample

<table>
<thead>
<tr>
<th>Group</th>
<th>C A&lt;sup&gt;a&lt;/sup&gt;</th>
<th>M A&lt;sup&gt;b&lt;/sup&gt;</th>
<th>I Q&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Read. Achieve.</th>
<th>Arith. Achieve.</th>
<th>Parent educ.</th>
<th>Effort rating</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G L&lt;sup&gt;d&lt;/sup&gt;</td>
<td>S S&lt;sup&gt;e&lt;/sup&gt;</td>
<td>G L&lt;sup&gt;f&lt;/sup&gt;</td>
<td>S S&lt;sup&gt;g&lt;/sup&gt;</td>
</tr>
<tr>
<td>A</td>
<td>mean</td>
<td>143.10</td>
<td>93.95</td>
<td>67.80</td>
<td>21.55</td>
<td>67.00</td>
<td>29.75</td>
</tr>
<tr>
<td></td>
<td>S.D.</td>
<td>14.33</td>
<td>14.15</td>
<td>7.17</td>
<td>8.96</td>
<td>5.28</td>
<td>10.17</td>
</tr>
<tr>
<td>B</td>
<td>mean</td>
<td>136.10</td>
<td>92.00</td>
<td>67.45</td>
<td>21.90</td>
<td>66.55</td>
<td>29.10</td>
</tr>
<tr>
<td></td>
<td>S.D.</td>
<td>12.40</td>
<td>14.12</td>
<td>8.35</td>
<td>11.73</td>
<td>8.29</td>
<td>12.17</td>
</tr>
<tr>
<td>C</td>
<td>mean</td>
<td>140.05</td>
<td>91.55</td>
<td>66.80</td>
<td>24.20</td>
<td>69.25</td>
<td>28.50</td>
</tr>
<tr>
<td></td>
<td>S.D.</td>
<td>15.67</td>
<td>15.31</td>
<td>7.32</td>
<td>7.16</td>
<td>5.40</td>
<td>9.08</td>
</tr>
<tr>
<td>D</td>
<td>mean</td>
<td>139.00</td>
<td>94.45</td>
<td>69.40</td>
<td>21.40</td>
<td>67.45</td>
<td>28.95</td>
</tr>
<tr>
<td></td>
<td>S.D.</td>
<td>14.73</td>
<td>13.97</td>
<td>6.92</td>
<td>6.46</td>
<td>5.05</td>
<td>7.88</td>
</tr>
<tr>
<td>Entire Sample</td>
<td>mean</td>
<td>139.56</td>
<td>92.99</td>
<td>67.86</td>
<td>22.26</td>
<td>67.56</td>
<td>29.08</td>
</tr>
</tbody>
</table>

<sup>a</sup> Chronological Age  
<sup>b</sup> Mental Age  
<sup>c</sup> Intelligence Quotient  
<sup>d</sup> Grade Level  
<sup>e</sup> Standard Score  
<sup>f</sup> Grade Level  
<sup>g</sup> Standard Score  
<sup>h</sup> Socioeconomic Index
inventory. After his preference was stated for an item, the examiner marked the proper space on the answer sheet. This method of administration was used to avoid error from poor reading skills or careless marking of the answer sheet. A copy of the answer sheet is included in Appendix D. It was designed by the writer to reveal clearly the total number of choices for each category (16 maximum) as well as the number of times each category was selected over each of the other categories. This type of data revealed the consistency of the subject's selections as well as the total number of his choices in each category. The data of concern for this study were each subject's most and least preferred categories.

The Reward Preference Inventory was given to each subject approximately one week before his performance on the experimental task. This time period corresponded with reliability data on the Reward Preference Inventory reported by Dunn-Rankin and Shimizu (1968). Exceptions to the time interval between the RPI and the experimental task were made for subjects for whom most and least preferred RPI categories had remained stable during a two month period. The RPI was administered at two month intervals during a concurrent study of the long-term stability of preference category ranks (Clum, 1970). Some of the same subjects were used in both studies. Identical RPI administration procedures were used in both studies.

Base Rate Task. One of the major shortcomings of Cartwright's study (1968) was an absence of base rate information about the amount of
work that might have been done by her subjects had none of the Dunn-Rankin rewards been offered as incentives. Without base rate information it would have been impossible to know whether the amount of effort a child expended to achieve an extrinsic reward would have been significantly less had no reward been offered. Therefore, each child in the educable mentally retarded classes from which the sample was selected was given an arithmetic task that would challenge his current skill level. Several different tasks were used. Samples of each are included in Appendix G. The particular task selected for each student was based upon his teacher's judgment of the type of problem that would challenge him. A challenging, school-like task was used to avoid the novelty effect that appeared to be present in Cartwright's tasks (1968). The effect of situational specificity and the dangers of overgeneralization from a laboratory to a reality situation have been discussed by Mischel (1968), Murstein (1965) and Rotter (1954).

The tasks were long enough to extend well beyond the usual interest span of most intermediate educable mentally retarded children on arithmetic tasks. The task length was designed to assure that the full extent of most children's intrinsic interest in arithmetic tasks would be measured. This was necessary so that reinforcement points on subsequent experimental tasks could be established well beyond each child's point of minimal fatigue or task boredom.

Base rate tasks were given to an entire class simultaneously.
The time selected for administration of the base rate task was after lunch (12:30 P.M.). An afternoon time was chosen since teachers claimed that children were less willing to do sustained work on academic tasks in afternoons than in mornings. An additional practical benefit would result if children's academic efforts during afternoons could be increased through extrinsic rewards.

The directions for the base rate task were worded carefully to avoid any suggestion that a certain number of problems should be worked. Each child was told to work as many problems as he desired, to try to answer correctly all that he worked, and to stop whenever he wished (see Appendix F). This was done to determine the number of problems each child would work without extrinsic incentives or expectation for a specific amount of work. The base rate directions were similar to Cartwright's experimental task directions in telling the child to stop working whenever he wished.

Each child's base rate task was scored for number of problems worked, percentage of problems correct, and amount of time worked. The time was recorded as accurately as possible; however, it was frequently necessary to record time to the nearest thirty seconds since several children often turned in their paper simultaneously.

Care was taken to determine that an appropriate task had been selected for each student, a task neither too hard nor too easy. If less than 60 percent of the child's problems were correct, an easier base rate task was given to him. If the entire set of
problems was completed with high accuracy in a short time a more
difficult base rate task was given to the child the following day.
If an adequate base rate performance was not achieved on the second
task, the child was not included in the sample.

Experimental Task. Each subject was given an experimental task,
determined individually from his base rate, preference and group
assignment. The experimental task was presented from one to two
weeks after the base rate task and within one week after the Reward
Preference Inventory (with exceptions on the RPI cited above). The
interval between base rate and experimental tasks was selected because
of their similarity. A longer interval (more than two weeks) might
have resulted in increased arithmetic skills that would make a
similar task easier. A shorter interval (less than one week) might
have resulted in enough recall of specific answers to make a similar
task easier. This second possibility was mitigated somewhat by the
number and similarity of the problems.

High similarity between base rate and experimental tasks was
needed to equate difficulty level between the tasks. Equal difficulty
level was needed so that task characteristics would not vary to an
extent that would confound the effects of the extrinsic incentives
offered during the experimental task. A control group was included
to assess the degree of task equality. Control group conditions
replicated those of the base rate task. The control group was
included because of three possible results of task similarity.
First, significantly more problems might be done the second time. This could be due to a practice effect, increased arithmetic skills or recall of specific answers from the base rate task. Second, significantly fewer problems might be done the second time. This could result from task boredom or discouragement from lack of extrinsic incentives. Third, a similar number of problems might be done the second time. This could result from "true" task equality, or a balance resulting from the factors mentioned above or other factors.

Comparison of control group with base rate results, then, would provide an indication of task equality.

The experimental task was given in the same classroom setting as the base rate task. It was given at the same time of day as the base rate task (12:30 P.M.). The directions for the experimental task, however, were given to each subject individually immediately preceding the task.

This was done outside the classroom so that no child's directions were overheard by others. Each child was asked to refrain from telling others about his directions or incentives. The child was shown his task; a red line indicated how many problems he would need to work to receive a reward from his most preferred Dunn-Rankin category; a green line indicated how many problems he would need to work to receive a reward from his least preferred category. The rewards in each of the categories were read to the child as he was shown the appropriate line. He was asked if the directions were clear to him or if he had
any questions about the task or the rewards. Questions were answered or directions clarified if necessary. (See Appendix F for directions to each of the groups).

Four different groups were established for the experimental task. Subjects in group A were required to work at least twenty-five percent beyond their (individual) base rate to select a reward from their least preferred Dunn-Rankin category or at least fifty percent beyond base rate to select a reward from their most preferred category. Only one reward was selected by each subject, i.e. a subject working fifty percent or more beyond his base rate could select a reward from his most preferred category only. Group A was included to answer question 3A, whether the value of a highly preferred reward (on the Dunn-Rankin) could overcome a child's disinclination to exert the effort needed to achieve it.

Group B subjects were required to work only 25 percent beyond base rate to choose a reward from their most preferred category, but fifty percent beyond base rate to choose from their least preferred category. Here, too, each subject selected only one reward. Group B was included to answer question 3B whether educable mentally retarded subjects would be controlled more by extrinsic reward value than by a drive to accomplish the most difficult task despite little extrinsic reward value.

Group C subjects were allowed to choose a reward from either their most or least preferred categories after working twenty-five
percent beyond base rate. Group C was included to answer questions 2, whether significantly more educable mentally retarded children would select a reward from their most preferred than from their least preferred category after working twenty-five percent beyond base rate.

Group D subjects received no reward regardless of how many problems they worked. Their experimental task was a replication of base rate task conditions except that directions were given individually. Group D was included to answer question 1 and questions about effects of task similarity raised above.

TABLE 3
Experimental Conditions for Each Group

<table>
<thead>
<tr>
<th>Group</th>
<th>Amount of effort beyond base rate required for reward</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25%</td>
</tr>
<tr>
<td>A</td>
<td>5</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>5 or 1\textsuperscript{a}</td>
</tr>
<tr>
<td>D</td>
<td>control - no reward</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Numbers indicate preference rank on the Dunn-Rankin Reward Preference Inventory.

Scoring of the experimental task was similar to that for the base rate task. Number of problems done, percentage of problems correct and time to the nearest thirty seconds was determined for each subject.
Selection and delivery of reward. After completion of the experimental task the papers were scanned to determine which subjects in the experimental groups (A B C) had completed enough problems to receive a reward, i.e. to or beyond a reinforcement point. Each child who had earned a reward was given a choice of one of the items in the appropriate category. Each child in groups A and B selected a reward from the category (his most or least preferred) appropriate to the reinforcement point that he had reached or passed. Each subject in group C that had worked at least twenty-five percent beyond his base rate was allowed to select a reward from either his most or least preferred category. The rewards were delivered within several days of selection. The delay between task performance and reward delivery may have biased the results against the research hypotheses since delayed reinforcement has been found to be less effective than immediate reinforcement.

Temporal sequence of methodology and procedures

The temporal sequence of methodology and procedures used with each classroom is summarized in Chart I.
CHART I

Flow Chart of Methodology and Procedures Used with Each Classroom

DAY NUMBER

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>7</th>
<th>8-13</th>
<th>14</th>
<th>15</th>
<th>16-17</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Base rate task**
- **Second base rate task for those needing it**
- **Random assignment to experimental groups**
- **Reward selection by eligible subjects**
- **Delivery of selected rewards to eligible subjects**

- **A** (LPR\(^a\) = 25%, MPR\(^b\) = 50%)
- **B** (MPR = 25%, LPR = 50%)
- **C** (MPR = 25%, LPR = 25%)
- **D** (No reward)

---

- **Base rate task**
- **Reward Preference Inventory**
- **Experimental task**

---

a Least preferred reward
b Most preferred reward
Analysis of Data. Question 1 asked whether subjects in the reinforcement groups (A B C) would perform a significantly greater amount of work beyond their base rate on a second administration of the academic task than would non-reinforced subjects. Question 1 was explored by establishing reinforcement points beyond each reinforced subject's base rate as indicated in Table 6. The question would be answered positively if the proportion of subjects in the reinforcement groups (A B C) working to a point at least 25% beyond their base rate was significantly greater than the proportion of subjects in the non-reinforcement group (d) working to a similar point. A 2x2 contingency table and the chi-square test for two independent samples (Siegel, 1956; pp. 104-110) were used to determine the significance of the results.

Question 2 asked whether upon reaching a reinforcement point beyond base rate on an academic master task a significantly greater number of subjects in group C would select a reward from their most preferred category than from their least preferred category on the Reward Preference Inventory. Question 2 was explored by establishing a reinforcement point 25% beyond each child's base rate. A choice from each child's most and least preferred categories was offered upon reaching this point. Consistent selection from the most preferred category was considered a task validation of the Dunn-Rankin Reward Preference Inventory. The Chi-square ($X^2$) one-sample test was used to determine the significance of the results (Siegel, 1956; pp. 42-47).
Question 3 asked whether a significantly greater number of subjects would work beyond their base rate on an academic mastery task to the point of their most preferred category than to the point of their least preferred category both when **more** effort were required to reach their most preferred category and when **less** effort were required. Question 3 was explored by establishing reinforcement points 25% and 50% beyond each child's base rate. In group A a child reached his least preferred category at the 25% point and his most preferred category at the 50% point. In group B each child reached his most preferred category at the 25% point and his least preferred category at the 50% point. Data gathered to answer this question were the number of subjects working to the point at which they would receive a reward from their most preferred category compared to the number of subjects working to the point at which they would receive a reward from their least preferred category. The significance of the results was determined for groups A and B combined and individually by the $X^2$ one-sample test (Siegel, 1956; pp. 42-47).

**Additional Analyses.** Intercorrelations among all descriptive independent and dependent variables were computed and the significance of each was determined. Of particular interest were relationships between achievement and perseverance measures. These included intercorrelations among standard score of arithmetic achievement on the Wide Range Achievement Test, effort ratings by teachers, and number of problems on the base rate task. Positive intercorrelations were expected among these variables since Cartwright (1968) found that...
academic achievement was related to length of performance on a per-
severance task.

Performance differences between base rate and experimental
tasks were compared among groups. The differences were used to
answer questions raised above (p.38) about the effects of repeating
base rate conditions with the control group.

A comparison was made between stated preferences on the Reward
Preference Inventory and rewards actually received as a result of
the amount of work done on the experimental task. Discrepancies
found through this comparison could suggest different values for
reward categories than revealed by stated preferences. However, care-
ful analysis of experimental conditions was done before conclusions
were stated.

Summary

Socioeconomic aspects of the school and community, methodology
and procedures of the study and data analysis techniques used were
described in this chapter.
CHAPTER III
Results and Discussion

Hypotheses of the study, results for the hypotheses, additional analyses of the data and discussion of results will be included in this chapter.

The study was designed to answer the questions posed above. From these questions the following hypotheses were derived:

I. Reinforcement preferences will control the amount of effort applied to arithmetic tasks by students in intermediate educable mentally retarded classes to a significantly greater extend (a. = .05) than minimal fatigue or task boredom.

II. The most preferred reward, as measured by the Reward Preference Inventory, of students in intermediate educable mentally retarded classes will be chosen significantly (a. = .05) more often than will the least preferred reward when a choice of either is offered upon completion of an arithmetic task that requires effort beyond the point of minimal fatigue or task boredom (base rate).

III. Significantly (a. = .05) more students in intermediate educable mentally retarded classes will work to the point beyond their base rate where they will receive their most preferred reward, as measured by the Reward Preference Inventory, than will work to the point where they receive their least preferred reward . . .

A. whether the point at which they receive their most preferred reward requires more effort to reach than the point at which they receive their least preferred reward or . . .
B. whether the point at which they receive their most preferred reward requires less effort to reach than the point at which they receive their least preferred reward.

Results

The general hypothesis was that preferred rewards (as indicated by the Dunn-Rankin Reward Preference Inventory) could control the amount of effort intermediate age students in educable mentally retarded classes would exert upon an academic task. Each research hypothesis was stated in null form and tested.

Null Hypothesis I

There will be no significant difference ($\alpha = .05$) between the proportion of subjects in the experimental and control groups working at least 25% beyond their base rate on the experimental task.

Results. The chi-square test for two independent samples was need to determine the significance of results for hypothesis I (Siegel, 1956, p. 104). Null hypothesis I was rejected since a chi-square ($\chi^2$) of 30.1 ($p < .001$) was obtained for the difference between the proportion of experimental and control subjects working at least 25% beyond base rate on the experimental task. Only three of 60 subjects in the experimental groups worked less than 25% beyond base rate on the experimental task while 13 of 20 control subjects worked less than 25% beyond their base rate. Only seven control subjects worked more than 25% beyond their base rate. These results are summarized in Table 4 below.
TABLE 4

Chi-Square Contingency Table for Hypothesis I

<table>
<thead>
<tr>
<th></th>
<th>Subjects working at least 25% beyond their base rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>YES</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Experimental groups (ABC)</td>
<td>57</td>
</tr>
<tr>
<td>Control group (D)</td>
<td>7</td>
</tr>
</tbody>
</table>

\[ X^2 = 30.1 \quad p < .001 \]

Results demonstrated that RPI rewards, when given for specific amounts of work, could increase the academic efforts of students in intermediate educable mentally retarded classes since 57 of 60 experimental subjects worked at least 25% beyond their base rate. This suggested some utility for RPI rewards in the classroom.

Subsidiary Analysis. A question raised in designing methodology to assess the RPI asked whether educable mentally retarded students would do more, less, or the same amount of work on a second administration of an academic task when no reward was offered. A control group was included. Since base rates alone could not answer this question. The difference in number and percentage of problems worked between base rate and experimental tasks was determined for each subject. Means and standard deviations in these areas were determined for experimental and control subjects. Table 5 summarizes the contrasts between experimental and control subject performance from base rate to experimental task.
TABLE 5

Means and Standard Deviations of Experimental and Control Subject Performance on Base Rate Task and Experimental Task

<table>
<thead>
<tr>
<th>Task performance</th>
<th>Experimental subjects</th>
<th>Control subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>Number of problems on BRT a</td>
<td>80.05</td>
<td>89.20</td>
</tr>
<tr>
<td>on ET b</td>
<td>118.53</td>
<td>94.20</td>
</tr>
<tr>
<td>Increase in number of problems from BRT to ET</td>
<td>38.48</td>
<td>5.00</td>
</tr>
<tr>
<td>Percentage of increase in problems from BRT to ET</td>
<td>43.07%</td>
<td>5.60%</td>
</tr>
</tbody>
</table>

a Base Rate Task
b Experimental Task
c Standard deviation

The mean percentage of increase in problems worked from base rate to experimental task was 43.1% for experimental subjects but only 5.6% for control subjects. Control subjects worked slightly more problems than experimental subjects on the base rate task while experimental subjects worked substantially more problems on the experimental task (see Table 5). Experimental subjects also had a substantially greater increase in number and percentage of problems worked from base rate task to experimental task. However, the large standard deviations in Table 5 reflected great individual variability among both experimental and control subjects. Several subjects, both
experimental and control, showed variations of over 100 problems between base rate and experimental tasks. The large and varying standard deviations and the unequal sample size precluded further statistical analysis of results (McNemar, 1962, p. 106).

Table 5 suggested answers to several important questions about RPI utility that were not answered by Cartwright's (1968) study. These were: "How will the performance of control subjects differ from base rate to experimental tasks?" "How will control subject performance differences contrast with similar differences among experimental subjects?" Control subject performance appears to differ very little from base rate to experimental task when group means are compared. By contrast experimental subjects' performance increases markedly from base rate to experimental task, again, when group means are compared.

Table 4 showed that seven control subjects worked more than 25% beyond their base rate. Three were in the same classroom and comprised all the control subjects from that class (class 5). Class 5 was the only classroom in which every student, both experimental and control, worked at least 25% beyond his base rate. A tentative inference was made that the task persistence of class 5 subjects was due to external pressure for achievement created by the teacher's behavior. The directions to the control group on both the base rate and experimental tasks were to stop working problems whenever the subject wanted to or became tired. These directions were similar to
those given by Cartwright (1968) to all her subjects. During the administration of the base rate and experimental tasks the teacher urged her class to "show Mr. _______ (the experimenter) how well you can do". These comments occurred despite a prior discussion with the teacher about the importance of avoiding any statements that would indicate an expectation for a certain amount of work.

Because of teacher exhortation to "do well" and other conditions discussed below, results for hypotheses II and III were computed for the entire sample and again with subjects from class 5 excluded. However, results for hypothesis I were not re-computed with class 5 subjects excluded since significance of results for the entire sample was beyond .01.

Hypothesis I conclusions

1. Hypothesis I results showed that rewards from the RFI could increase the efforts of a significant number of subjects in intermediate age classes for educable mentally retarded students on repetitive arithmetic tasks at least 25% beyond their base rate when effort-reward contingencies were not offered.

2. Results suggested also that exhortation to "do well" by the class 5 teacher may have increased academic effort 25% beyond base rate.
3. The utility of the control group in revealing different trends of results among the classroom suggested that control groups and repeated base rate measures be retained in the methodology of future studies.

**Null Hypothesis II**

There will be no significant difference (α = .05) between the number of subjects selecting a reward from their most preferred or their least preferred category on the Reward Preference Inventory when such a choice is offered each subject who works at least 25% beyond his base rate on an arithmetic task.

**Results.** Nineteen of 20 subjects in group C met the reward criterion by working at least 25% beyond their base rate on the experimental task. Thirteen subjects selected a reward from their most preferred category while six selected a reward from their least preferred category. This proportion resulted in a chi-square (1,19) of 2.58 on the chi-square one sample test. A chi-square of 2.71 was required to reach the .05 level of significance for a one-tailed test. Therefore, null hypothesis II was not rejected for the entire sample. Table 6 summarized the results for null hypothesis II.
TABLE 6
Chi-Square Contingency Table for Hypothesis II
Entire Sample

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Preference category of reward selection</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Most</td>
<td>Least</td>
</tr>
<tr>
<td>Observed</td>
<td>13</td>
<td>6.</td>
</tr>
<tr>
<td>Expected</td>
<td>9.5</td>
<td>9.5</td>
</tr>
</tbody>
</table>

Chi-square 2.58\(^a\)  \(p > .05\)

\(^a\) Chi-square of 2.71 required for significance at .05 level (one-tailed test)

Although null hypothesis II could not be rejected for the entire sample, the results were near significance in the predicted direction. A class by class analysis of the results revealed that three of the six subjects that selected a least preferred reward were from class 5. Two of these three subjects selected a consumable reward (their least preferred category) instead of adult approval or peer reward (their most preferred category). Further analysis of results for all subjects in class 5 indicated that eight of 12 subjects eligible to receive a reward selected a least preferred category although six of these eight subjects could have selected a most preferred category for the same amount or less effort on the experimental task. Seven of the eight students indicated consumable rewards (\(^\oplus\) um, pop, ice cream, candy bar) were their least preferred category.
Selection of consumable rewards, despite stated preference for adult approval and peer rewards, might have resulted from student awareness of teacher difficulty in dispensing social rewards. This impression was derived from informal observation of teacher behavior in the classroom. When told of the most preferred reward category of several subjects, the teacher did not believe their selections. Also, the teacher did not demonstrate skills in delivering such rewards (peer, competitive, adult approval) in the classroom. Experimenter suggestions concerning delivery of these rewards in the classroom did not change teacher behaviors. In light of the effects of expectancy upon behavior reported by Rotter (1954) it seemed quite possible that several students in class 5, aware of the teacher's difficulty in dispensing social rewards might have chosen to receive a less preferred, but certain, consumable reward from the experimenter than to risk not receiving a highly preferred social reward requiring teacher implementation.

Subsidiary Analysis. Hypothesis II results were re-computed with class 5 subjects eliminated. Because hypothesis II was important to RPI validity, because the chi-square results came very close to rejecting the null hypothesis, and because class 5 conditions confounded the effects of reward value with expectancy.

Twelve of the 15 remaining subjects in group C selected a most preferred reward while three selected a least preferred reward. This proportion resulted in a chi-square (1,15) of 5.4 on the chi-square
one sample test. A chi-square of 5.41 was required to reach the .01 level of significance for a one-tailed test; however, a chi-square of only 2.71 was required to reach the .05 level of significance. Therefore, null hypothesis II was rejected when class 5 results were eliminated. Table 7 summarized the results for null hypothesis II with class 5 results eliminated.

**TABLE 7**

Chi-Square Contingency Table for Hypothesis II
Class 5 Results Eliminated

<table>
<thead>
<tr>
<th>Frequency (Group C subjects)</th>
<th>Preference category of reward selection</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Most</td>
<td>Observed: 12</td>
<td>3</td>
</tr>
<tr>
<td>Least</td>
<td>Expected: 7.5</td>
<td>7.5</td>
</tr>
<tr>
<td></td>
<td>N: 15</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Chi-square 5.4</strong></td>
<td><strong>p &lt; .05</strong></td>
</tr>
</tbody>
</table>

Hypothesis II conclusions

1. Rejection of null hypothesis II with class 5 results eliminated demonstrated that stated preferences on the Reward Preference Inventory could predict actual reward selections following an academic task when an equal amount of effort beyond base rate was required to select a reward from either most or least preferred reward categories. Some empirical support for the task validity of category preferences on the RPI was provided by these results.
2. Class 5 results demonstrated that possible effects of expectancy upon reward selection may exceed the effects of individual preference. Through their selection of consumable rewards, despite low preference for them, class 5 subjects expressed reservations about their teacher's willingness to mediate social rewards in the classroom.

3. A question remained about why three subjects (not in class 5) had selected a least preferred reward when a most preferred reward was available for equal effort. This suggested that factors in addition to possible expectancy and stated preferences could influence reward selection. Information about such factors was sought through additional analyses of results (reported below).

Null Hypothesis III

There will be no significant difference (a. = .05) between the number of subjects working beyond their base rate on an arithmetic task to a point where they will receive a reward from their most preferred category and the number working to a point where they will receive a reward from their least preferred category . . .

A. whether the point at which they receive their most preferred reward requires more effort to reach than the point at which they receive their least preferred reward or . . .

B. whether the point at which they receive their most preferred reward requires less effort to reach than the point at which they receive their least preferred reward.

Results. Forty subjects were selected to test hypothesis III, twenty each in group A and group B. Each subject in group A was required to work 25% beyond his base rate to select a least preferred
reward or 50% beyond his base rate to select a most preferred reward. The reward conditions were reversed for subjects in group B. The initial analysis of results for hypothesis III combined results for groups A and B. Thirty eight subjects (the total from group A and B eligible to receive a reward by working at least 25% beyond base rate) were included in the initial analysis. They were divided according to whether they had worked to the point of eligibility for a reward from their most preferred category or from their least preferred category.

Eleven subjects in group A became eligible for a most preferred reward by working at least 50% beyond base rate. Eight group A subjects were eligible for a least preferred reward by working between 25% and 50% beyond base rate. One group A subject worked less than 25% beyond base rate and was not eligible for a reward. Twelve subjects in group B became eligible for a reward from their most preferred category by working between 25% and 50% beyond base rate. Seven group B subjects became eligible for a least preferred reward by working at least 50% beyond base rate. One group B subject worked less than 25% beyond base rate and was not eligible for a reward.

The chi-square one sample test determined the significance of results for the 38 subjects in groups A and B who had worked at least 25% beyond their base rate on the experimental task. A chi-square (1,38) of 1.68 was obtained which, although beyond the .10 level in the predicted direction, was inadequate to reject null
hypothesis III. Results for the initial analysis of groups A and B combined are summarized in Table 8.

### TABLE 8

Chi-Square Contingency Table for Hypothesis III Results
All Subjects Eligible for Reward

<table>
<thead>
<tr>
<th>Frequency (Groups A &amp; B)</th>
<th>Number of subjects selecting reward category</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Most preferred</td>
<td>Least preferred</td>
</tr>
<tr>
<td>Observed</td>
<td>23</td>
<td>15</td>
</tr>
<tr>
<td>Expected</td>
<td>19</td>
<td>19</td>
</tr>
</tbody>
</table>

Chi-square 1.68

a Chi-square of 2.71 needed for significance at .05 level (one-tailed test)

Although the initial analysis of groups A and B combined failed to reach an acceptable level of significance (a. = .05) for rejection of Null Hypothesis III, further analysis was needed because of the effects of expectancy and teacher behavior upon class 5 subjects reported above. Therefore, a subsidiary analysis of results for hypothesis III was done with class 5 results excluded.

**Subsidiary Analysis.** Thirty subjects remained after class 5 results were eliminated, fifteen each in group A and group B. Nine subjects in group A were eligible for a most preferred reward while six subjects were eligible for a least preferred reward. Eleven group B subjects were eligible for a most preferred reward while four subjects...
were eligible for a least preferred reward. The chi-square one sample test determined the significance of the subsidiary analysis. A chi-square (1,30) of 3.34 indicated significance beyond the .05 level in the predicted direction. Null hypothesis III could be rejected if results for class 5 were eliminated. Subsidiary analysis results for hypothesis III are summarized in Table 9.

**TABLE 9**

Chi-Square Contingency Table for Hypothesis III
Class 5 Results Eliminated

<table>
<thead>
<tr>
<th>Frequency (Groups A &amp; B)</th>
<th>Subjects selecting reward category</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Most preferred</td>
<td>Least preferred</td>
</tr>
<tr>
<td>Observed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>Expected</td>
<td>15</td>
<td>15</td>
</tr>
</tbody>
</table>

| Chi-square 3.34 | p < .05 |

Rejection of null hypothesis III demonstrates that a most preferred reward category, as determined by the RPI could control the amount of effort expended upon an academic task significantly more often than a least preferred reward category. The effects of expectancy upon academic performance were again demonstrated through class 5 results. Class 5 students again behaved as though their teacher could not implement preferred social rewards. The discrepancy between results for class 5 students and the other experimental subjects (groups A B C) seem to support Rotter's Social Learning Theory (1954) since it demonstrated that reinforcement value could
control academic performance when expectancy was controlled, but not when expectancy varied among available rewards.

Although null hypothesis III was rejected when results for groups A and B were combined, conclusions about the classroom utility of the RPI required that the results for each group be analyzed separately since group A conditions required that subjects work 50% beyond base rate for their most preferred reward while group B subjects could receive their most preferred reward by working only 25% beyond base rate. Hypothesis IIIA (group A) predicted that a reward from a most preferred RPI category could evoke 25% more academic effort (50% beyond base rate) than a reward from a least preferred RPI category. Hypothesis IIIB (group B) predicted that a most preferred RPI category, when available at a point 25% beyond base rate, could stop academic effort short of a point 50% beyond base rate at which a least preferred RPI category reward was available. Hypothesis IIIB was included to determine whether students in group A who had worked 50% beyond base rate for a most preferred reward did so because of reward value or because of a "need to achieve", i.e. compulsion to attain the most difficult goal available regardless of reward value.

Subsidiary Analysis - Hypotheses IIIA + IIIB. Group A was used to analyze results for hypothesis IIIA while group B was used to analyze results for hypothesis IIIB. Because of conditions described above, the data were analyzed with results for subjects from class 5 eliminated. The chi-square one sample test was used for the separate
analyses of group A and group B. Nine subjects from group A were eligible for a most preferred reward. Six subjects were eligible for a least preferred reward. A chi-square (1,15) of .6 was obtained which was not adequate to reject null hypothesis IIIA. (Table 10).

Eleven subjects from group B were eligible for a most preferred reward. Four subjects were eligible for a least preferred reward. A chi-square (1,15) of 3.26 was significant beyond the .05 level in the predicted direction which was adequate to reject null hypothesis IIIB. Results of separate analyses of hypotheses IIIA and IIIB are summarized in Table 10.

TABLE 10
Summary of Separate Analyses of Hypotheses IIIA and IIIB
Class 5 Data Eliminated

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Subjects selecting reward category</th>
<th>Chi-square</th>
<th>Significance level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Most preferred</td>
<td>Least preferred</td>
<td></td>
</tr>
<tr>
<td>III A</td>
<td>9</td>
<td>6</td>
<td>.6^a</td>
</tr>
<tr>
<td>III B</td>
<td>11</td>
<td>4</td>
<td>3.26</td>
</tr>
</tbody>
</table>

^a Chi-square of 2.71 needed for significance at .05 level (one-tailed test)

Results for hypothesis IIIA revealed that a most preferred RPI category failed to evoke effort 50% beyond base rate (25% beyond the point of eligibility for a least preferred reward) in a significant number of subjects. However, the results were in the predicted direction since more subjects worked 50% beyond base rate (nine)
than 25% beyond base rate (six). Results for hypothesis IIIB demonstrated that a most preferred RPI category could prevent a significant number of subjects from working 25% beyond the point of availability of their most preferred RPI category. One possible interpretation of hypothesis IIIB results was that reward preferences had influenced academic performance to a greater extent than a "need to achieve" the most difficult goal available.

Comparison of group A and group B results suggested that a most preferred reward could evoke academic effort 50% beyond base rate more often than a least preferred reward. A chi-square test for two independent samples was done to determine whether a most preferred RPI category had evoked academic effort 50% beyond base rate significantly more often than a least preferred RPI category. Results for the analysis (Table 11) did not reach significance, but revealed a trend (p < .10) in the expected direction. This suggested that a most preferred RPI category could evoke academic effort 50% beyond base rate more often than a least preferred RPI category.
Chi-Square Contingency Table of Results of Comparative Ability of Most and Least Preferred RPI Categories to Evoke Effort 50% Beyond Base Rate
Class 5 Results Eliminated

<table>
<thead>
<tr>
<th>Group</th>
<th>Subjects working beyond base rate</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25%</td>
<td>50%</td>
</tr>
<tr>
<td>A</td>
<td>6 (LPR)\textsuperscript{a}</td>
<td>9 (MPR)</td>
</tr>
<tr>
<td>B</td>
<td>11 (MPR)\textsuperscript{b}</td>
<td>4 (LPR)</td>
</tr>
</tbody>
</table>

Chi-square 2.17\textsuperscript{c}  \( p < .10 \)

\textsuperscript{a} Least Preferred Reward

\textsuperscript{b} Most Preferred Reward

\textsuperscript{c} Chi-square of 2.71 needed for significance at .05 level (one-tailed test)

Hypothesis III conclusions

1. A most preferred RPI category controlled the academic efforts of students in intermediate age educable mentally retarded classes significantly (\( a. = .05 \)) more often than a least preferred RPI category.

2. A most preferred RPI category failed to increase academic effort 50% beyond base rate when a least preferred RPI category was available 25% beyond base rate. However, results suggested that a most preferred RPI category might increase academic effort more than 25% beyond base rate, but less than 50% beyond base rate when a least preferred category was available 25% beyond base rate.
3. A most preferred RPI category, available 25% beyond base rate, prevented academic effort 50% beyond base rate for receipt of a least preferred RPI category. From these results it was inferred that reward preference affected academic effort to a greater extent than "need for achievement".

4. Expectancy level for receipt of a RPI reward might affect academic performance to a greater extent than reward preference.

**Summary of cumulative experimental results.** A summary of cumulative results for all experimental subjects on hypotheses II and III determined the need for further analyses. Sixty subjects were divided equally among three experimental groups (A B C). Twenty others were in a control group (D). Fifty-seven of the 60 experimental subjects (19 in each experimental group) worked at least 25% beyond their base rate on the experimental task and became eligible to receive a reward. Thirty-six of the 57 subjects performed, as predicted, by selecting a most preferred reward after working to criterion. Twenty-one of the 57 subjects failed to perform as predicted by selecting a least preferred reward after working to criterion. Seventeen of the 21 subjects who failed to perform as predicted selected a consumable reward. The remaining four subjects selected a peer reward. The summary above suggests differential predictability and strength among the categories since consumable and peer rewards were selected only when least preferred. Conversely, adult approval, competitive and independence rewards were selected only when most preferred.
Eight of the 21 subjects who failed to perform as predicted were in class 5. Observations of teacher behavior before and during administration of base rate and experimental tasks to class 5 subjects suggested that differential expectancy of reward categories may influence academic efforts more than stated reward value. A less plausible explanation was that teacher exhortation to "do well" had driven class 5 subjects so far on the base rate task that they were incapable of 50% more effort on the experimental task. The latter explanation seems less plausible than the differential expectancy explanation because three of four class 5 subjects in group B worked 50% beyond base rate for a least preferred consumable reward.

Another six of the 21 subjects who failed to perform as predicted were in group A. A plausible explanation for these results was that the extra effort required of group A subjects for receipt of a most preferred reward (25%) exceeded the difference in value between most and least preferred RPI categories for the six group A subjects. Although plausible, further analysis seemed necessary before accepting this explanation since nine of 15 group A subjects had worked 50% beyond base rate for a most preferred reward while seven subjects form groups B and C selected a reward from their least preferred RPI category. The seven group B and C subjects could have selected a reward from their most preferred RPI category for equal or less effort. These results suggested that functions other than preference statements affected reward value. Further analyses were
Additional Analyses

Summary of Reward Preference Inventory Choices

Mean and standard deviation of choices in each reward category for each group and entire sample are summarized in Table 12. Adult approval rewards were most preferred while consumable rewards were least preferred. However, Table 13 shows that only 37 of 80 subjects actually selected adult approval rewards most frequently while only 40 of 80 subjects chose consumable rewards least often. Therefore, if group preferences been used as rewards for each subject, the selection would have been inaccurate for more than 50% of the subjects. These data were consistent with Cartwright's (1968) concern about the need to individualize rewards.
<table>
<thead>
<tr>
<th>Groups</th>
<th>Adult Approval</th>
<th>C&lt;sup&gt;a&lt;/sup&gt;</th>
<th>C&lt;sup&gt;b&lt;/sup&gt;</th>
<th>I&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Peer</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Mean</td>
<td>11.20</td>
<td>8.40</td>
<td>4.65</td>
<td>6.85</td>
<td>8.95</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>3.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B Mean</td>
<td>10.15</td>
<td>8.60</td>
<td>6.00</td>
<td>7.05</td>
<td>8.20</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>3.03</td>
<td>3.35</td>
<td>5.16</td>
<td>3.53</td>
<td>3.49</td>
</tr>
<tr>
<td>C Mean</td>
<td>10.15</td>
<td>9.15</td>
<td>4.45</td>
<td>7.05</td>
<td>9.25</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>2.72</td>
<td>2.18</td>
<td>3.72</td>
<td>3.27</td>
<td>3.91</td>
</tr>
<tr>
<td>D Mean</td>
<td>12.30</td>
<td>8.05</td>
<td>5.30</td>
<td>6.45</td>
<td>7.85</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>3.08</td>
<td>3.30</td>
<td>3.79</td>
<td>3.05</td>
<td>2.48</td>
</tr>
<tr>
<td>Entire Sample Mean</td>
<td>10.95</td>
<td>8.55</td>
<td>5.10</td>
<td>6.85</td>
<td>8.56</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>3.06</td>
<td>2.98</td>
<td>4.23</td>
<td>3.04</td>
<td>3.56</td>
</tr>
</tbody>
</table>

<sup>a</sup> Competition
<sup>b</sup> Consumable rewards
<sup>c</sup> Independence
**TABLE 13**

Frequency of Most and Least Preferred Reward Categories for Each Group and the Entire Sample

<table>
<thead>
<tr>
<th>Groups</th>
<th>Most preferred</th>
<th>Least preferred</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AA (^a)</td>
<td>C (^b)</td>
</tr>
<tr>
<td>A</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>C</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>D</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>37</td>
<td>7</td>
</tr>
</tbody>
</table>

\(^a\) Adult approval  
\(^b\) Competition  
\(^c\) Consumable rewards  
\(^d\) Independence

Fifty-seven experimental subjects worked enough problems on the experimental task to reach criteria for reward. The number of subjects selecting a reward from each category is summarized in Table 14.
TABLE 14

Frequency of Reward Category Selection
By Subjects Eligible for a Reward

<table>
<thead>
<tr>
<th>Groups</th>
<th>Adult approval</th>
<th>Competition</th>
<th>Consumable rewards</th>
<th>Independence</th>
<th>Peer</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>5</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>1</td>
<td>10</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>C</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>12</td>
<td>3</td>
<td>25</td>
<td>2</td>
<td>15</td>
</tr>
</tbody>
</table>

Although Table 13 indicated that adult approval rewards were most frequently preferred on the RPI, Table 14 revealed that consumable rewards were most frequently selected after the experimental task. Further analysis was done to explore this discrepancy and to identify possible reasons for it.

**Differential effects of the Reward Preference Inventory**

*Importance.* The major focus of this study was the impact of reward value upon academic performance. Reward value was determined through preference statements on the RPI. Therefore, the question was raised, "How frequently did the subjects work to the point on the experimental task where they became eligible for their most preferred RPI category?" The answer to this question reflected the predictability of the RPI. Predictability was the most important characteristic of the RPI as a measure of reward value.
The predictability of the RPI within each experimental group was explored by the hypotheses. However, predictability of the RPI within each classroom was not explored. Predictability of each RPI category within each experimental group and classroom also was not explored. Differential predictability by category, by experimental group and by classroom might reveal patterns suggesting that academic performance was affected by factors other than reward value, expectancy and effort requirements. If indicators of other factor effects appear, the design or procedures of future studies might be altered to control them.

**Predictability.** A procedure was devised for determination of predictability to allow comparison among categories, experimental groups and classrooms. The procedure could be used either with the total RPI or with categories from the RPI. The procedure involved dividing the number of times the RPI (or category) predicted correctly within an experimental group or classroom by the number of times the RPI (or category) made a prediction for that experimental group or classroom. The quotient (expressed as a %) controlled for differential frequency of prediction among categories, experimental groups and classrooms.

Predictability was determined for the RPI for all experimental subjects, each experimental group and each classroom. The predictability of each RPI category was determined also for all experimental subjects, experimental groups and classrooms. Results for all
experimental subjects and experimental groups were computed a second time with class 5 data eliminated. The results of the analyses of predictability are summarized in Table 15.

**TABLE 15**

Predictability of the Reward Preference Inventory and of Each RPI Category among Experimental Groups and Classrooms

<table>
<thead>
<tr>
<th>Subjects</th>
<th>RPI</th>
<th>C</th>
<th>R</th>
<th>Peer</th>
<th>AA</th>
<th>I</th>
<th>Competition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groups A,B,C</td>
<td>63a</td>
<td>52</td>
<td>68</td>
<td>58</td>
<td>81</td>
<td>78</td>
<td>78</td>
</tr>
<tr>
<td>minus class 5</td>
<td>71</td>
<td>62</td>
<td>82</td>
<td>62</td>
<td>86</td>
<td>78</td>
<td>78</td>
</tr>
<tr>
<td>Group A</td>
<td>58</td>
<td>53</td>
<td>57</td>
<td>60</td>
<td>67</td>
<td>67</td>
<td>67</td>
</tr>
<tr>
<td>minus class 5</td>
<td>60</td>
<td>55</td>
<td>75</td>
<td>56</td>
<td>67</td>
<td>67</td>
<td>67</td>
</tr>
<tr>
<td>Group B</td>
<td>63</td>
<td>46</td>
<td>71</td>
<td>57</td>
<td>85</td>
<td>67</td>
<td>67</td>
</tr>
<tr>
<td>minus class 5</td>
<td>73</td>
<td>60</td>
<td>80</td>
<td>67</td>
<td>100</td>
<td>67</td>
<td>67</td>
</tr>
<tr>
<td>Group C</td>
<td>68</td>
<td>58</td>
<td>73</td>
<td>57</td>
<td>80</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>minus class 5</td>
<td>80</td>
<td>75</td>
<td>85</td>
<td>67</td>
<td>80</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Class 1</td>
<td>92</td>
<td>100</td>
<td>80</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>75</td>
</tr>
<tr>
<td>Class 2</td>
<td>67</td>
<td>60</td>
<td>100</td>
<td>20</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Class 3</td>
<td>40</td>
<td>17</td>
<td>50</td>
<td>20</td>
<td>75</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Class 4</td>
<td>82</td>
<td>78</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Class 5</td>
<td>33</td>
<td>27</td>
<td>38</td>
<td>33</td>
<td>50</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a Numbers indicate percentage of time the RPI (or category) predicted academic performance accurately

b Consumable rewards
c Adult approval
d Independence
The most obvious pattern was the low predictability of all RPI categories for class 5 subjects and the increased predictability for experimental groups with class 5 data removed. The class 5 pattern was attributed to the expectancy effects discussed above.

Another pattern of predictability among classrooms included an inverse relationship between predictability and age in classes 1-3 with a sudden rise in predictability in class 4. (Classes were numbered according to chronological age with class 1 as the youngest.) The pattern was clear only the consumable reward and adult approval categories, class 5 effects dominated results in the peer category. The independence and competitive reward categories had little influence upon predictability patterns because of their infrequent appearance as either most or least preferred RPI categories. The low predictability for the consumable and adult approval categories in classes 2 and 3 could have resulted from three possible conditions.

1. A category was the highest preference on the RPI more frequently than it was selected following the academic task.

2. A category was selected following the academic task more frequently than it was the highest preference on the RPI.

3. A category's preference and selection discrepancy approached zero through nearly equal occurrence of conditions 1 and 2.

To facilitate discussion of conditions 1-3 the following terminology was used.

**Predicted strength** - The number of times a category was most
preferred on the RPI divided by the frequency of the category.

**Frequency**  - The total number of times a category was either most or least preferred on the RPI.

**Actual strength**  - The number of times a category was selected following the academic task divided by the frequency of the category.

**Discrepancy**  - The difference between predicted and actual strength for a category. The discrepancy was preceded by a minus sign when predicted strength was greater than actual strength and by a plus sign when predicted strength was less than actual strength.

The adult approval and consumable reward categories had a lower predictability than other categories in the experimental groups as well as in classes 2 and 3. Therefore, analyses of frequency, predicted strength, actual strength and discrepancy were done for the adult approval and consumable reward categories in each experimental group and each classroom to determine which conditions described above (1-3) led to lower predictability. Results of the analyses are summarized in Table 16.
### TABLE 16

Strength and Discrepancy Analyses for Consumable Reward and Adult Approval Categories

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Consumable rewards</th>
<th>Adult approval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Freq.</td>
<td>PS</td>
</tr>
<tr>
<td>Groups</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A,B,C</td>
<td>40^e</td>
<td>25^f</td>
</tr>
<tr>
<td>Group A</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>Group B</td>
<td>13</td>
<td>38</td>
</tr>
<tr>
<td>Group C</td>
<td>12</td>
<td>25</td>
</tr>
<tr>
<td>Class 1</td>
<td>4</td>
<td>50</td>
</tr>
<tr>
<td>Class 2</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>Class 3</td>
<td>6</td>
<td>33</td>
</tr>
<tr>
<td>Class 4</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Class 5</td>
<td>11</td>
<td>9</td>
</tr>
</tbody>
</table>

- Frequency
- Predicted strength
- Actual strength
- Discrepancy
- Number of times available
- Percentage of frequency most preferred
- Percentage of frequency selected after task
- Difference between predicted and actual strength
Several patterns emerged from Table 16 that suggested effects upon reward value in addition to individual preference statements. The most obvious pattern included a consistent predicted strength less than actual strength discrepancy in the consumable reward category and a consistent predicted strength greater than actual strength discrepancy in the adult approval category. The discrepancy pattern occurred in all experimental groups and in most classrooms. The only classroom exceptions included predicted strength equal to actual strength for both categories in class 1 and predicted strength equal to actual strength for the adult approval category in class 4. Large discrepancies for both categories occurred in classes 2, 3, and 5; small or no discrepancies occurred in classes 1 and 4.

Table 16 revealed that predicted strength for the consumable reward category decreased with age while predicted strength for the adult approval category increased with age (class 1- youngest, class 5- oldest). This pattern was consistent with previous RPI results (Dunn-Rankin and Shimizu, 1968). A similar pattern occurred for actual strength in the consumable reward and adult approval categories, but on two distinct levels – one for classes 1 and 4, another for classes 2, 3 and 5.

The bi-level pattern of the consumable reward and adult approval category discrepancies (great in classes 2, 3 and 5; low or nil in classes 1 and 4) suggested the operation of differential effects
among the classrooms. Novelty appeared to be an effect of the categories among the classrooms. Adult approval rewards (grades) are used daily in nearly every classroom. Most teachers never use consumable rewards for academic effort in the classroom. A few teachers use consumable rewards infrequently such as on Friday afternoon if class behavior and work has been "good" during the week. Consumable rewards are seldom used systematically for a specified amount of effort on a task, and the novelty of consumable reward availability for a specified amount of effort could have enhanced its value for many subjects (Kish, 1966). Some evidence for the novelty interpretation was provided by analysis of predicted strength, actual strength and discrepancy for consumable and adult approval rewards in classrooms. Only one classroom (class 4) made systematic use of consumable rewards for academic effort. Class 4 had the lowest predicted and actual strength for the consumable reward category as well as the second lowest discrepancy between predicted and actual strength. Class 4 had the highest predicted strength (shared with two other classes) and the highest actual strength (alone) for the adult approval category. Class 4 with another class demonstrated the lowest discrepancy (0) between predicted and actual strength for the adult approval category. These results suggest that a portion of unpredicted consumable reward strength and adult approval weakness may be due to novelty of reward categories. Table 17 summarizes results of the predicted strength-actual strength comparison in all classrooms.
### TABLE 17
Comparison of Predicted Strength and Actual Strength of Consumable Reward and Adult Approval Categories in all Classrooms

<table>
<thead>
<tr>
<th>Classrooms</th>
<th>Consumable rewards</th>
<th>Adult approval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P S&lt;sup&gt;a&lt;/sup&gt;</td>
<td>A S&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Youngest</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 1</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>Class 2</td>
<td>50%</td>
<td>90%</td>
</tr>
<tr>
<td>Class 3</td>
<td>33%</td>
<td>83%</td>
</tr>
<tr>
<td>Class 4&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0%</td>
<td>22%</td>
</tr>
<tr>
<td>Oldest</td>
<td>Class 5</td>
<td>9%</td>
</tr>
</tbody>
</table>

<sup>a</sup> Predicted strength  
<sup>b</sup> Actual strength  
<sup>c</sup> Discrepancy  
<sup>d</sup> Classroom using consumable rewards systematically

Another effect upon the bi-level discrepancy pattern in the consumable reward and adult approval categories appeared to result from an interaction among individual maturity level, subtle reinforcement by adults of behaviors that reflect adult values, "social desirability" and RPI administration techniques. Teachers and parents frequently reinforce child behaviors that approximate adult values. Child behaviors that do not approximate adult values are sometimes rejected. This differential reinforcement may have greater effect with increasing age and social sensitivity of the child. Until puberty children may try to meet expressed or implied adult expectations.
These functions may result in greater conformity of the child's verbal behavior to adult expectations along with an increasing discrepancy between the child's verbal and actual behavior.

The RPI administration techniques used in the study involved face to face verbal interaction between each subject and an adult. This technique was used to minimize reading and marking errors on the RPI. The adult student interaction during RPI administration could have produced student verbal responses that teachers and parents had reinforced in the past. This may have resulted in verbal responses of greater maturity (high adult approval, low consumable reward) but less consistent with actual behavior (reward selection following academic task). These effects, combined with the possible influence of differential novelty among the categories, could have led to the bi-level discrepancy pattern illustrated in Table 16.

Another factor leading to zero discrepancies for class 1 in both categories could have been classroom control techniques used by the teacher. The class 1 teacher relied upon direct orders and punishment contingencies rather than upon exhortation or subtle reinforcement of "mature" behavior. Maturity level (class 1 was youngest) and the absence of subtle reinforcement of "mature" behavior might have given class 1 subjects less reason to feign sophisticated reward values than older subjects.
Summary and Implications. Additional analyses of category patterns were done to identify possible effects upon reward value or experimental results needing further exploration. No conclusions could be drawn from the data because of low frequencies and inadequate controls.

Additional analyses revealed a bi-level discrepancy pattern in the consumable reward and adult approval categories along with expected maturity trends (decreasing consumable reward increasing adult approval). Several effects were hypothesized for different aspects of the pattern. Low predicted strength - actual strength discrepancies in class 4 were attributed to a novelty effect since varied rewards and effort-reward contingencies were typical only in class 4. Low predicted strength - actual strength discrepancies in class 1 were attributed to a combination of maturity and socialization effects including adult pressures toward maturity, increasing conformity to adult values, teacher methods of classroom control and RPI administration techniques.

Clarification is needed of the influence of novelty and maturity-socialization effects upon reward value and academic performance. Effects of novelty could be assessed through comparison of RPI predictability and predicted strength - actual strength discrepancy patterns in classrooms regularly using a variety of rewards and effort-reward contingencies and in traditional classrooms. Results of the current study would suggest higher RPI predictability and lower predicted strength - actual strength discrepancy in classrooms.
using varied reward-effort contingencies.

Maturity-socialization effects could be assessed through comparison of RPI predictability and the predicted strength of RPI categories among different RPI administration techniques involving varied levels of impersonality. Such a comparison might include three groups with a different RPI administration procedure for each group. Group 1 could receive individual administration of the RPI as in the current study. Group 2 could receive a group administration of the RPI carefully structured to insure reading and marking accuracy with the results of each subject identified. Group 3 could receive the same controlled group administration procedure as group 2 with each subject's results anonymous. Results of the current study would suggest a positive relationship between impersonality of administration technique and RPI predictability, a positive relationship between impersonality and immaturity of preference (predicted strength higher for consumable reward, lower for adult approval), and an inverse relationship between impersonality and predicted strength – actual strength discrepancy.

The bi-level discrepancy patterns in the adult approval and consumable reward categories might be construed to attribute expectancy effects operating in class 5 to classes 2 and 3, also. The inclusion of class 5 with classes 2 and 3 in the bi-level pattern would suggest a basis for this interpretation. However, this potential interpretation can be eliminated by a summary of results
for the peer category in each classroom and by comparison of class 5 and non-class 5 results in each experimental group. Table 18 revealed a large predicted strength greater than actual strength discrepancy for class 5 peer results and either no discrepancy or substantial predicted strength less than actual strength discrepancies in classes 1-4. Table 18 revealed also predicted strength less than actual strength discrepancies for non-class 5 subjects in all experimental groups and either no discrepancy (group C) or a sizeable predicted strength greater than actual strength discrepancy (groups A and B) for class 5 subjects in the experimental groups. These patterns suggest unique effects in class 5 that are different from classes 2 and 3.
### TABLE 18
Strength and Discrepancy Analyses for the Peer Category

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Frequency</th>
<th>Predicted strength</th>
<th>Actual strength</th>
<th>Discrepancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groups A, B, C</td>
<td>25(^a)</td>
<td>60(^b)</td>
<td>60(^c)</td>
<td>0(^d)</td>
</tr>
<tr>
<td>Group A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>classes 1-4</td>
<td>4</td>
<td>75</td>
<td>100</td>
<td>+25</td>
</tr>
<tr>
<td>class 5</td>
<td>3</td>
<td>100</td>
<td>33</td>
<td>-67</td>
</tr>
<tr>
<td>Group B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>classes 1-4</td>
<td>5</td>
<td>20</td>
<td>40</td>
<td>+20</td>
</tr>
<tr>
<td>class 5</td>
<td>2</td>
<td>100</td>
<td>50</td>
<td>-50</td>
</tr>
<tr>
<td>Group C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>classes 1-4</td>
<td>8</td>
<td>50</td>
<td>63</td>
<td>+13</td>
</tr>
<tr>
<td>class 5</td>
<td>3</td>
<td>67</td>
<td>67</td>
<td>0</td>
</tr>
<tr>
<td>Class 1</td>
<td>5</td>
<td>40</td>
<td>60</td>
<td>+20</td>
</tr>
<tr>
<td>Class 2</td>
<td>3</td>
<td>33</td>
<td>33</td>
<td>0</td>
</tr>
<tr>
<td>Class 3</td>
<td>4</td>
<td>25</td>
<td>75</td>
<td>+50</td>
</tr>
<tr>
<td>Class 4</td>
<td>5</td>
<td>80</td>
<td>80</td>
<td>0</td>
</tr>
<tr>
<td>Class 5</td>
<td>8</td>
<td>88</td>
<td>50</td>
<td>-38</td>
</tr>
</tbody>
</table>

\(^a\) Number of times available  
\(^b\) Percentage of frequency most preferred  
\(^c\) Percentage of frequency selected after task  
\(^d\) Difference between predicted and actual strength

The competitive and independence categories were selected too infrequently for meaningful analyses of patterns among classrooms or experimental groups.

Analyses of category strength were done to locate discrepancy patterns among experimental groups and classrooms, to identify effects that might have lead to the discrepancy patterns, and to
suggest methods of exploration or control of the effects. Novelty and maturity-sociability were effects that seemed to influence reward value in addition to individual preference statements. If these effects can be controlled in future studies, RPI predictability should increase.

**Intercorrelation Table**

Intercorrelations were computed among all independent, dependent and descriptive variables and significance was determined (Bruning and Kintz, 1968). The question raised above about the relationship between achievement and perseverance found by Cartwright (1968) was investigated. Cartwright found a significant positive correlation between achievement and standard score for fives on her perseverance task (1968, p. 53). Support for Cartwright's findings in the current study included significant (.05 level) positive correlations between teacher ratings of effort on arithmetic tasks and standard score of arithmetic achievement, number of problems on the base rate task and number of problems on the experimental task. However, there were no significant positive correlations between standard score of arithmetic achievement and number of problems on base rate and experimental task. Therefore, only partial support was found in the current study for Cartwright's achievement-perseverance relationship. The intercorrelations are summarized in Table 19.
Several patterns of significant negative correlation were found among categories of the RPI. The consumable reward category had significant ($p < .001$) negative correlations with the adult approval, competition and peer reward categories. The independance reward category was similar but with a differing level of significance for each category, i.e. adult approval .001, competition .01, peer .05. There were no significant positive correlations between any of the categories. A similar pattern of bipolarities emerged from a study of 87 fourth, fifth and sixth grade students in the University of Hawaii Elementary School (Dunn-Rankin and Shimizu, 1968). The mean Intelligence Quotient of the students from Hawaii was 121 while the mean Intelligence Quotient of the 80 subjects in the current study was 68. The similar bipolarity patterns in both studies suggested
consistent clusters of preferences across ability levels. The intercorrelations among RPI categories are summarized in Table 20.

**TABLE 20**

<table>
<thead>
<tr>
<th>Intercorrelations among Reward Preference Inventory Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult approval</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Competition</td>
</tr>
<tr>
<td>Consumable rewards</td>
</tr>
<tr>
<td>Independence</td>
</tr>
<tr>
<td>Peer</td>
</tr>
</tbody>
</table>

\(a\) .36 required for significance at the .001 level  
\(b\) .28 required for significance at the .01 level  
\(c\) .22 required for significance at the .05 level

**Discussion**

Three topics will be considered in this section; the strengths and limitations of the design and procedures of the study, the implications for classroom use of the RPI, and the relationship between statements of individual reward preference and reward utility in school situations.

**Strengths and Limitations of Design and Procedures**

This section attempts to answer the question, "How well did the design and procedures of the study assess the classroom utility of the Reward Preference Inventory?" One approach to answering this question
is to consider the extent to which the current study was able to overcome the shortcomings of Cartwright's design and methodology.

A major deficiency in Cartwright's methodology was a lack of contingency between task effort and reward. The current study established criterion levels for reward at least 25% beyond the point at which each subject had stopped work on the task when no reward was offered (base rate). This was done to determine whether RPI rewards had enough value to evoke effort on the task beyond the subject's point of minimal fatigue or task boredom. The absence of an effort-reward contingency in Cartwright's study left in doubt the question of whether RPI rewards could evoke effort increments meaningful to a classroom teacher.

A second deficiency in Cartwright's methodology was that the rewards offered to the subjects did not correspond to the items in the RPI categories. The current study allowed each subject to select one of the four items from the reward category that he had earned on the experimental task. Each subject was informed which items he would be allowed to choose during his experimental task directions.

The third deficiency in Cartwright's study involved experimental tasks that little resemblance to school task. The current study used pages of arithmetic problems at a difficulty level appropriate to each subject. The appropriateness of the task was based upon performance criteria (described above) related to accuracy, speed and number of problems completed. These steps were taken to make the
experimental task similar to a classroom drill assignment in arithmetic.

The fourth deficiency in Cartwright's study involved evidence that subjects were too old or sophisticated to value the RPI rewards. Subjects in the current study were students in intermediate age classes for the educable mentally retarded. Both groups of subjects (Cartwright's and the current study) had a similar chronological age range. Subjects in the current study had a lower mental age range that may have been more appropriate for RPI rewards. No subject in the current study refused to select or accept a reward from the category earned through his effort on the experimental task. Fifty-seven of 60 experimental subjects worked at least 25% beyond their base rate to select a reward from an RPI category. These results suggested that the RPI rewards did have value to the experimental subjects.

The fifth deficiency in Cartwright's study was the lack of base rate data about how much effort might have been evoked from each subject on the experimental task had none of the RPI rewards been offered. The current study included a base rate task to determine how far each subject would work without RPI rewards offered. Reward points were then set 25% or 50% beyond each experimental subject's base rate on a highly similar experimental task to determine reward efficacy.
Another approach to answering the question, "How well did the design and procedures of the study assess the classroom utility of the Reward Preference Inventory?", is to consider, methodology used to explore each major question of the study and to evaluate how definitively the methodology has answered each question.

1. Will reward preferences, indicated by the Reward Preference Inventory, cause school children to work beyond their point of minimal fatigue or boredom (base rate) on an academic task?

Question 1 was answered definitively by a combination of base rate task, control group and effort-reward contingencies. The base rate task determined each subject's point of minimal fatigue or boredom on the academic task. The control group clarified the effect of base rate task repetition upon academic output. The effort-reward contingencies determined whether RPI rewards could evoke a "significant" (25%-50%) increase in effort on an academic task.

2. Will an intermediate age educably mentally retarded child's most preferred reward, as determined by the Reward Preference Inventory, be chosen consistently over his least preferred reward when a choice of either reward is offered upon completion of an academic task that requires effort beyond his base rate?

Question 2 was answered definitively by a design that required all subjects to work 25% beyond their base rate on an academic task to select a reward from either their most or least preferred RPI category. This design determined whether a significantly greater number of subjects would select a reward from their most preferred
RPI category than would select a reward from their least preferred RPI category when equal effort beyond base rate was required for either reward.

3. Will a most preferred reward control the effort of a student in an intermediate age educable mentally retarded class on an academic task significantly more often than a least preferred reward regardless of whether his most preferred reward requires . . .

A. more effort or . . .

B. less effort beyond base rate than his least preferred reward?

Question 3 was answered less definitively than questions 1 or 2 because the amount of effort required for reinforcement was set arbitrarily at 25% and 50% beyond base rate. A better design would have retained the effort-reward contingency, but allowed the effort discrepancy between most and least preferred rewards to vary. A specific suggestion for such a design will be included below in the implications for future research.

A third approach to the "design and procedures" question should specify the limitations and inadequancies of the methodology of the study and suggest ways to overcome them in future research. Several limitations and inadequancies of the current study are discussed below.

The RPI administration procedures used in the study might have confounded reward value by social desirability and maturity-socialization effects. Adult and subject contact during RPI administration
could have lead some subjects, more with increasing age, to state preferences for adult approval rewards, while actually preferring consumable rewards. Future studies might avoid similar confounding factors through use of group administration procedures. Group administration of the RPI might decrease verbal responses designed to please the adult tester rather than to reflect the actual reward values of the child.

The design and procedures for question 3 were only partially adequate since a general question about the comparative degree of effort control by most or least preferred RPI categories was tied to specific amounts of effort beyond base rate. As stated above, a more adequate design for question 3 would have retained the effort-reward contingency while allowing variation in the effort discrepancy between most and least preferred rewards.

The base rate methodology assessed the amount of effort each subject would put forth on an academic task when no RPI rewards were offered. Also, care was taken by the experimenter to avoid emitting expectations for a certain amount of effort on the base rate task. These procedures were used to determine the ability of RPI rewards to evoke academic effort in situations devoid of teacher exhortation or offers of reward. Realistic school situations usually includes exhortations or rewards such as those brought to bare upon class 5 subjects (All class 5 control subjects worked at least 25% beyond base rate). Therefore, a comparison between RPI rewards and a
"standard" school situation would have been more relevant to the determination of the incremental utility of the RPI to evoke academic effort. The difficulty in determining such incremental utility was the construction of a "standard treatment" base rate situation that could have been justifiably considered "standard" for schools. However, such a comparison is needed to determine the value of any new method compared with "standard" school situations (Gardner, 1969). A future study might include a "standard treatment" base rate or control group along with an "absence of external pressure" group to determine the realistic utility of the RPI.

Despite a relatively homogeneous population, wide individual variability occurred in the study. The methodology provided no acceptable means to control or explore individual variability. Four control subjects (not in class 5) worked 25% or more beyond base rate on the experimental task for no reward either offered or delivered. Several possible reasons for this performance could be offered; however, continued base rate measures, as in a standard behavior modification study, would help determine when academic performance stabilizes. Individual variability might be handled better in future studies by repeated base rate measures.

Implications for classroom utility of the Reward Preference Inventory

Results for the hypotheses indicated that RPI rewards promised utility in controlling academic effort. This section will review and discuss some positive and negative indicators for the RPI in
control of academic effort.

**Positive Indicators of Utility.** Hypothesis I revealed that the RPI rewards evoked academic effort 25% beyond base rate in 57 of 60 experimental subjects. Hypothesis II showed that a most preferred RPI category was selected significantly (a. = .05) more often than a least preferred category when both were available for equal effort. Hypothesis III demonstrated that most preferred rewards were more effective than least preferred rewards in controlling effort on an academic task. Control was most effective when the most preferred RPI category required less academic effort than the least preferred category. Control of effort by a most preferred RPI category was greater than that of a least preferred category when the most preferred category required more effort, but the results for the comparison (Hypothesis IIIA) did not reach significance. (a. = .05) Therefore, every hypothesis suggested or demonstrated that a most preferred RPI category could control academic effort better than a least preferred RPI category. The major question that remained unanswered by the results was "How much additional academic effort will a most preferred RPI category evoke than a least preferred category?" Future studies should be designed to explore this question more specifically. One possible design will be described in the implications for future research.

**Negative Indicators of Utility.** Wide individual variability of results despite a relatively homogeneous population indicated limited
RPI predictability in individual instances. The wide variability might have been controlled more effectively through a design that included repeated base rate measures.

The possible influence of extraneous factors upon results (novelty, RPI administration techniques, maturity-socialization effects) suggested that individual preference statements alone were not an adequate indication of reward value. Cartwright's study (1968) and the current study were based upon the assumption that they were. Further study of the interaction of individual preference statements and the extraneous factors mentioned above should be conducted to determine the effects of these interactions upon academic effort.

Hypothesis IIIA results suggested that no RPI rewards might be powerful enough to evoke effort 50% beyond base rate. This could be a plausible interpretation of Hypothesis IIIA results because rewards of individual adult attention were not included on the RPI. Much research has shown that individual, positive attention from an adult is a powerful reward for most children (Broden, et.al., 1970; Schutte and Hopkins, 1970). The exclusion of this type of reward from the RPI might have been a factor in the failure of hypothesis IIIA to reach significance and might hamper severely the potential classroom utility of the RPI.

The RPI included no assessment of the differential effectiveness
of individuals dispensing the rewards. Tharp and Wetzel (1969) have suggested that characteristics of reward dispensers may have a significant effect upon reward value. This limitation, too, could hamper the potential classroom utility of the RPI.

Preference Statements and Reward Utility. The assumption made by Cartwright (1968) and by the current study that statements of individual preference would outweigh the importance of other factors affecting reward value received little support from results of the current study. Novelty, RPI administration techniques and maturity-socialization effects appear to interact with individual preference statements to determine reward utility. Characteristics of reward dispensers also appear to influence reward utility. Research is needed to explore the interactions of the effects listed above and their influence upon reward efficacy in school classrooms.
CHAPTER IV
SUMMARY AND CONCLUSIONS

Chapter IV summarizes the importance of the problem, the questions selected for investigation and the procedures used to answer the questions. The conclusions and the limitations of the study along with suggestions for further research are included in Chapter IV also.

Importance of the Problem

Effort has been found to be a key factor in academic achievement. Individual differences in reward preference have been found to affect the amount of effort expended upon a performance task. Statements of preference have been a popular method used to identify effective rewards for individuals. The Reward Preference Inventory (RPI) has been the most comprehensive, systematic and efficient instrument reported to assess individual differences in reward preference among children in school classrooms. Only one study (Cartwright, 1968) has been reported of the utility of the RPI in increasing the amount of effort school children expend upon academic tasks. Cartwright's study contained several methodological deficiencies that made it inadequate as an evaluation of the RPI.
Reward preference inventories are easy to construct and convenient to administer. They appear likely to become a popular method to assess reinforcement preferences. As more teachers become aware of these factors, reward preference inventories seem destined to proliferate. Therefore, methodology and procedures to evaluate the classroom utility of reward preference inventories need to be developed before proliferation occurs.

Questions

The study attempted to assess the classroom utility of the RPI; however, several aspects of the RPI had to be investigated prior to classroom utility. These included whether RPI rewards were valued highly enough by the subjects to evoke academic effort beyond the point of minimal fatigue or boredom, and whether school children actually valued a reward from a most preferred RPI category more than a reward from a least preferred category. Therefore, the questions posed for the study asked whether RPI rewards could evoke effort on an academic task beyond the point of minimal fatigue or boredom, whether a reward from a most preferred RPI category would be selected over a reward from an least preferred RPI category when equal effort was required and whether a reward from a most preferred RPI category could control academic effort to a greater extent than a reward from a least preferred RPI category.

An additional aspect of the study included assessment of the predictability of each of the five RPI categories among experimental
groups and classrooms.

Procedures

Eighty subjects from intermediate age classes for educable mentally retarded students were used to limit intellectual variability, to avoid volunteer bias and to provide an appropriate maturity level for RPI rewards. Individual base rate measures were taken for each subject on a repetitious arithmetic task. No extrinsic rewards were offered for base rate performance. Several tasks were used so that each subject's task would be appropriate to his arithmetic skill level.

The RPI was administered to each subject individually one week before experimental task performance. Most and least preferred RPI categories were determined for each subject.

The eighty subjects were assigned randomly to four groups. Group A subjects could select a least preferred RPI reward for experimental task effort 25% beyond (individual) base rate. They could select a most preferred RPI reward for effort 50% beyond base rate. Group B subjects could select a least preferred RPI reward for effort 50% beyond base rate and a most preferred RPI reward for effort 25% beyond base rate. Group C subjects could select either a most or least preferred RPI reward for effort 25% beyond base rate. Group D subjects (control group) were offered no rewards.

An experimental task was designed for each subject that was
similar to his base rate task. Tasks were individualized for each subject by base rate measure, RPI preferences and experimental group assignment. Experimental task directions were designed to clarify each subject's effort-reward contingency and to provide him with criterion achievement feedback during performance. Rewards were delivered to subjects earning them several days after experimental task performance.

Procedures for the study were designed to assess RPI utility through effort-reward contingencies beyond each subject's base rate. Rewards offered corresponded to RPI category statements. Task novelty was minimized through use of arithmetic drill assignments. Appropriate subject maturity level for RPI rewards was sought through use of students in intermediate age educable mentally retarded classes. The procedures attempted to overcome deficiencies noted in a prior assessment of RPI classroom utility.

Conclusions

1. RPI rewards were able to evoke significantly (a.= .05) more effort than task repetition with no RPI rewards.

2. Individual perference statements on the RPI predicted reward selection accurately when most and least preferred RPI categories were available for equal effort.

3. Most preferred RPI rewards appeared to control academic effort better than least preferred rewards. Limitations in the design
of the study precluded information about the extent to which control by a most preferred reward exceeded control by a least preferred reward.

4. Great individual variability of results suggested that preference statements alone are inadequate indicators of reward utility. Novelty, administration techniques, mediator preferences and maturity-socialization effects appear to influence reward utility also.

5. Potential classroom utility of the RPI may have been diminished by several omissions. No RPI category assessed preferences for individual adult attention. No RPI category included immediate knowledge of results. The RPI neglected to assess the effects of mediator value upon reward value.

Limitations of the study

1. Effort-reward criteria set arbitrarily 25% and 50% beyond base rate made hypothesis III methodology an inadequate indicator of whether more effort on an academic task could be evoked by a most preferred reward than by a least preferred reward. An adequate design would provide variability in the discrepancy between the amount of effort required for a most preferred reward and a least preferred reward.

2. The absence of control or base rate measures that reflected "standard" classroom conditions limited design adequacy as a measure
of RPI classroom utility. The control and base rate measures used compared performance when RPI rewards were offered with performance when they were not.

3. Use of a single base rate measure, rather than repeated base rate measures, may have lead to wide individual variability that confounded interpretation of results.

4. Face to face contact between examiner and subject during RPI administration may have contaminated preference with "social desirability" effects. Future studies should avoid potential contamination by using group administration procedures.

Suggestions for future research

1. Results of the current study suggested that hypothesis IIIA might have been confirmed had a smaller effort differential been required between most and least preferred RPI categories. A study is needed of the amount of additional academic effort that a most preferred RPI category could evoke beyond a least preferred RPI category. A possible design for the study should include several reward intervals beyond base rate between most and least preferred RPI categories.

2. A "standard procedure" control group is needed in addition to a "no reward" control group to demonstrate the incremental utility of RPI rewards in comparison with current classroom practices.
3. Future studies with the RPI should assign randomly subjects to individual and group administration procedures to assess the impact of face to face contact between examiner and subject upon preference statements.

4. Repeated base rate measures should be included in future studies to assess the extent of individual variability.

5. Modification of the RPI or development of new reward preference inventories should include assessment of preferences for individual adult attention, immediate knowledge of results and various reward mediators.

6. A study is needed of the effects of novelty upon RPI reward utility. The study should compare RPI predictability in classrooms using a variety of rewards and effort-reward contingencies with RPI predictability in traditional classrooms.

7. RPI studies should be extended to various populations and academic tasks to determine generality of effects.
APPENDIXES
APPENDIX A

Development of the Reward Preference Inventory

In 1965 a list of reinforcements that could theoretically be used by teachers in the elementary school classroom was formed. These rewards were classified into five categories: adult approval, competitive approval, peer approval, independence rewards, and consumable rewards. Instruments were constructed, based on the method of paired comparisons which scaled the reward preferences of school children in the fourth and sixth grades of the University Elementary School located in Tallahassee, Florida. Preference profiles were obtained for each child and correlations between reward preference, ability, sex, achievement, and socio-economic status were obtained as well as group profiles for fourth and sixth grade classes.

The results of this pilot study indicated that children differ widely as to which rewards they say they prefer and in the relative emphasis they place on the different categories of reward. Generally, however, adult approval was indicated as being most liked and consumable rewards as least preferred. Depending upon the sub-population, reward preference profiles correlated significantly with achievement and ability.

Results of this earlier work were encouraging enough to continue efforts to construct a more valid measure of relative reward preference. To increase the content validity of the original instrument, eighty additional statements reflecting reinforcers actually observed in the third through the sixth grades in the University Elementary School at the University of Hawaii were collected and combined with the original set which had been theoretically proposed. In order to obtain items in each preference category which were homogeneous in terms of value preference the statements were scaled on a (4 category) likert scale (Edwards, 1957) whose descriptions ranged from dislike to like very much.

Seventy students in the fourth, fifth, and sixth grades at the University Elementary School responded to this rating instrument during the summer of 1966 and plots of items in each category were made in which the abscissa contained discrimination indices.
and the ordinate item score values. Those items which plotted high in score value and low in discrimination were selected to represent the reward preference category. This procedure was an attempt to obtain "well liked" items in each reward preference category so that comparisons across categories would not be affected by differentially valued items. It was also important to obtain items which had low discrimination indices so that responses to statements within a category would be homogeneous over a large population of youngsters. Although a "coke" was well liked by the majority of the students it also yielded a high discrimination index and was therefore eliminated as a consumable reward, for example.

Subtests were constructed by drawing high valued but low discriminating items from each of five categories of rewards. All possible pairs were formed among these statements and presented according to the order procedures developed by Ross (1937). Paired comparison scaling (Dunn-Rankin, 1965) was used because it insures that all comparisons are made and because a dichotomous choice is usually an easier task than ranking for young children. Six separate subsets of ten item pairs were thus formed. These subtests were presented with directions as indicated in Appendix A.

This revised instrument was administered to 28 first grade children and 65 fourth grade children in the rural schools in Hawaii and to 18 Ponapean children in the sixth grade at Kolonia, Ponape, Micronesia. Analysis of the results of this administration revealed the same diversity of patterns originally observed and some interesting changes in the group profiles over grade level and cultural background.

The subtest profiles were scrutinized and where differences between category scores consistently exceeded two rank units the original statements were re-examined to see if such variability was a product of their ambiguity. This analysis suggested that some improvement could be made in the instrument so that each subtest profile was a more similar measure of relative reward preference.

A third revision of four subtests, each consisting of 10 independent pairings of five categorical statements according to the pairing procedures developed by Ross (1937) was made. (Dunn–Rankin and Shimizu, 1968, pp. 3-5)
APPENDIX B

Reward Preference Inventory

Name ___________________________________
Grade ___________________________________

Directions

Suppose that you have worked hard on an assignment and you think that you have done a good job. Which one of two things below would you most like to have happen?

A. _____ Teacher gives you a gold star.
B. _____ A friend says he likes you.

If you choose the gold star, mark an X on the line (____) in front of the statement about the gold star. If you like what your friend would say best, mark an X on the (____) in front of that sentence.

On the next two pages are pairs of other things that might happen after you have done a good job or finished your work. Choose one of each pair that you would like best and mark an X on the line (____) that comes in front of it. Mark only one line for each pair. Do you understand? Are there any questions?
1. A. _____ Teacher writes "100" on your paper.
   B. _____ Be first to finish your work.
2. A. _____ A package of bubble gum.
   B. _____ Students ask you to be on their team.
3. A. _____ Be free to do what you like.
   B. _____ Teacher writes "100" on your paper.
4. A. _____ Students ask you to be on their team.
   B. _____ Be first to finish your work.
5. A. _____ Be free to do what you like.
   B. _____ A package of bubble gum.
6. A. _____ Teacher writes "100" on your paper.
   B. _____ Students ask you to be on their team.
7. A. _____ Be first to finish your work.
   B. _____ Be free to do what you like.
8. A. _____ A package of bubble gum.
   B. _____ Teacher writes "100" on your paper.
9. A. _____ Students ask you to be on their team.
   B. _____ Be free to do what you like.
10. A. _____ Be first to finish your work.
    B. _____ A package of bubble gum.
11. A. ____ Teacher writes "A" on your paper.
    B. ____ Be the only one that can answer a question.

12. A. ____ A candy bar.
    B. ____ Friends ask you to sit with them.

13. A. ____ Be free to go outside.
    B. ____ Teacher writes "A" on your paper.

14. A. ____ Friends ask you to sit with them.
    B. ____ Be the only one that can answer a question.

15. A. ____ Be free to go outside.
    B. ____ A candy bar.

16. A. ____ Teacher writes "A" on your paper.
    B. ____ Friends ask you to sit with them.

17. A. ____ Be the only one that can answer a question.
    B. ____ Be free to go outside.

18. A. ____ A candy bar.
    B. ____ Teacher writes "A" on your paper.

19. A. ____ Friends ask you to sit with them.
    B. ____ Be free to go outside.

20. A. ____ Be the only one that can answer a question.
    B. ____ A candy bar.
21. A. _____ Teacher writes "Perfect" on your paper.

B. _____ Have only your paper shown to the class.

22. A. _____ An ice cream cone.

B. _____ Classmates ask you to be class leader.

23. A. _____ Be free to play outside.

B. _____ Teacher writes "Perfect" on your paper.

24. A. _____ Classmates ask you to be class leader.

B. _____ Have only your paper shown to the class.

25. A. _____ Be free to play outside.

B. _____ An ice cream cone.

26. A. _____ Teacher writes "Perfect" on your paper.

B. _____ Classmates ask you to be class leader.

27. A. _____ Have only your paper shown to the class.

B. _____ Be free to play outside.

28. A. _____ An ice cream cone.

B. _____ Teacher writes "Perfect" on your paper.

29. A. _____ Classmates ask you to be class leader.

B. _____ Be free to play outside.

30. A. _____ Have only your paper shown to the class.

B. _____ An ice cream cone.
31. A. _____ Teacher writes "Excellent" on your paper.
B. _____ Have your paper pot on the bulletin board.

32. A. _____ A soft drink.
B. _____ Friends ask you to work with them.

33. A. _____ Be free to work on something you like.
B. _____ Teacher writes "Excellent" on your paper.

34. A. _____ Friends ask you to work with them.
B. _____ Have your paper pot on the bulletin board.

35. A. _____ Be free to work on something you like.
B. _____ A soft drink.

36. A. _____ Teacher writes "Excellent" on your paper.
B. _____ Friends ask you to work with them.

37. A. _____ Have your paper put on the bulletin board.
B. _____ Be free to work on something you like.

38. A. _____ A soft drink.
B. _____ Teacher writes "Excellent" on your paper.

39. A. _____ Friends ask you to work with them.
B. _____ Be free to work on something you like.

40. A. _____ Have your paper pot on the bulletin board.
B. _____ A soft drink.
APPENDIX C

Reward Preference Inventory Categories and Rewards

Adult Approval
1. Teacher writes "100" on your paper.
2. Teacher writes "A" on your paper.
3. Teacher writes "perfect" on your paper.
4. Teacher writes "excellent" on your paper.

Competitive Approval
1. Be first to finish your work.
2. Be the only one that can answer a question.
3. Have only your paper shown to the class.
4. Have your paper put on the bulletin board.

Peer Approval
1. Students ask you to be on their team.
2. Friends ask you to sit with them.
3. Classmates ask you to be class leader.
4. Friends ask you to work with them.

Independence
1. Be free to do what you like.
2. Be free to go outside.
3. Be free to play outside.
4. Be free to work on something you like.

124
Consumable Rewards

1. A package of bubble gum.
2. A candy bar.
3. An ice cream cone.
APPENDIX D

Reward Preference Inventory

Answer Sheet

Name ___________________________ Grade ____ C.A._____ Date ______

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td></td>
<td>11.</td>
<td></td>
<td>21.</td>
<td></td>
<td>31.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td>12.</td>
<td></td>
<td>22.</td>
<td></td>
<td>32.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td></td>
<td>13.</td>
<td></td>
<td>23.</td>
<td></td>
<td>33.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td>14.</td>
<td></td>
<td>24.</td>
<td></td>
<td>34.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td></td>
<td>15.</td>
<td></td>
<td>25.</td>
<td></td>
<td>35.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td></td>
<td>16.</td>
<td></td>
<td>26.</td>
<td></td>
<td>36.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td></td>
<td>17.</td>
<td></td>
<td>27.</td>
<td></td>
<td>37.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td></td>
<td>18.</td>
<td></td>
<td>28.</td>
<td></td>
<td>38.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td></td>
<td>20.</td>
<td></td>
<td>30.</td>
<td></td>
<td>40.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>AA</th>
<th>C</th>
<th>CR</th>
<th>I</th>
<th>P</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Ranks

1. _____
2. _____
3. _____
4. _____
5. _____
APPENDIX E

Effort Ratings

Directions

Please rate each of the students in your class from A to E as to the fifth of your class in which they fall on the effort they put into arithmetic tasks. The letter ratings are to be used as follows:

- A - highest 20%
- B - second highest 20%
- C - middle 20%
- D - second lowest 20%
- E - lowest 20%

Each child should be rated on his standing in your class, not in terms of his individual potential or in comparison with regular students of his age.

Thank you for your help.
Base Rate Task

I have some arithmetic problems here. I would like you to do some of them for me. Do as many of them as you like. Try to get right most of those you do. Let me know when you get tired and want to stop. Are there any questions?

Hypothesis II, Group C

I have some more arithmetic problems here. If you do all of the problems to this line (pointing)\(^1\), you will get one of these. (Show child list of rewards from his fifth preference category and from his first preference category. Allow child time to read the reward list. Answer any questions he may have concerning the rewards on either list). You may have only one. If you see one you would like, do all of the problems up to this line (pointing) and you will get it. Try to get as many right as you can. Are there any questions?

Hypothesis III, Groups A and B

I have some more arithmetic problems here. If you do all of the problems to this green line (pointing)\(^2\), you will get one of these.

---

\(^1\) Line was 25% beyond base rate for each subject in Group C.

\(^2\) Green line was 25% beyond base rate for each subject in Group A and 50% beyond base rate for each subject in Group B.
(Show child list of rewards from his fifth preference category. Allow him time to read the reward list. Answer any questions he has concerning the rewards). If you do all of the problems to the red line (pointing)\(^3\), you will get one of these. (Show child list of rewards from his first preference category. Allow him time to read the list. Answer any questions he has concerning the rewards). You may have only one. If you see one you would like, do all of the problems up to the line for that one and you will get it. Try to get as many right as you can. Are there any questions?

**Control Group (D)**

I have some more arithmetic problems here. I would like you to do some of them for me. Do as many of them as you like. Try to get right most of those you do. Let me know when you get tired and want to stop. Are there any questions?

---

\(^3\) Red line was 50% beyond base rate for each subject in Group A and 25% beyond base rate for each subject in Group B.
APPENDIX G
Sample of Tasks

Numbers

<table>
<thead>
<tr>
<th>Name</th>
<th>Score</th>
<th>Preferences</th>
<th>Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grade</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>7 9 5 7 9 3 7</th>
<th>3 2</th>
<th>7 2 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2 0 9 7 8 3</td>
<td>9 8</td>
<td>9 1 6</td>
</tr>
<tr>
<td>6</td>
<td>6 8 1 4 4 3</td>
<td>1 1</td>
<td>3 2 2</td>
</tr>
<tr>
<td>9</td>
<td>5 5 0 4 3 6</td>
<td>0 3</td>
<td>5 0 7</td>
</tr>
<tr>
<td>6</td>
<td>1 7 1 8 6 9</td>
<td>3 8</td>
<td>1 2 8</td>
</tr>
<tr>
<td>0</td>
<td>4 7 8 1 3 1</td>
<td>8 1</td>
<td>7 2 6</td>
</tr>
<tr>
<td>3</td>
<td>9 2 0 1 0 8</td>
<td>1 0</td>
<td>6 3 6</td>
</tr>
<tr>
<td>9</td>
<td>9 8 6 5 2 6</td>
<td>4 4</td>
<td>2 0 0</td>
</tr>
<tr>
<td>8</td>
<td>5 9 5 7 0 1</td>
<td>4 7</td>
<td>2 5 7</td>
</tr>
<tr>
<td>8</td>
<td>5 4 7 4 9 7</td>
<td>1 9</td>
<td>2 8 8</td>
</tr>
<tr>
<td>1</td>
<td>7 1 9 5 4 9</td>
<td>3 1</td>
<td>0 1 8</td>
</tr>
<tr>
<td>5</td>
<td>3 7 3 2 4 0</td>
<td>4 9</td>
<td>4 0 5</td>
</tr>
<tr>
<td>5</td>
<td>5 4 4 3 3 9</td>
<td>1 1</td>
<td>0 1 6</td>
</tr>
<tr>
<td>8</td>
<td>0 8 4 2 9 3</td>
<td>7 9</td>
<td>4 8 2</td>
</tr>
<tr>
<td>4</td>
<td>4 8 6 3 0 6</td>
<td>1 7</td>
<td>3 1 0</td>
</tr>
<tr>
<td>7</td>
<td>7 9 5 6 8 7</td>
<td>2 1</td>
<td>4 0 8</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
## Addition Facts

Name ____________________________

Grade ____________________________

Age ____________________________

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>+2</td>
<td>+5</td>
<td>+2</td>
<td>+0</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>+3</td>
<td>+4</td>
<td>+6</td>
<td>+1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>+1</td>
<td>+4</td>
<td>+5</td>
<td>+7</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>+8</td>
<td>+3</td>
<td>+4</td>
<td>+9</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>+0</td>
<td>+8</td>
<td>+7</td>
<td>+4</td>
</tr>
</tbody>
</table>

Score ____________________________

Preferences ______________________

Percentages ______________________

132
### Subtraction Facts

<table>
<thead>
<tr>
<th>Name</th>
<th>Grade</th>
<th>Age</th>
<th>Score</th>
<th>Preferences</th>
<th>Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>-3</td>
<td>-7</td>
<td>-5</td>
<td>-8</td>
<td>-4</td>
<td>-2</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>9</td>
<td>12</td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td>-8</td>
<td>-9</td>
<td>-5</td>
<td>-7</td>
<td>-6</td>
<td>-2</td>
</tr>
<tr>
<td>14</td>
<td>11</td>
<td>8</td>
<td>10</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td>-0</td>
<td>-2</td>
<td>-6</td>
<td>-3</td>
<td>-4</td>
<td>-7</td>
</tr>
<tr>
<td>13</td>
<td>11</td>
<td>6</td>
<td>12</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td>-7</td>
<td>-4</td>
<td>-2</td>
<td>-5</td>
<td>-9</td>
<td>-8</td>
</tr>
<tr>
<td>16</td>
<td>5</td>
<td>12</td>
<td>9</td>
<td>17</td>
<td>10</td>
</tr>
<tr>
<td>-9</td>
<td>-2</td>
<td>-3</td>
<td>-6</td>
<td>-8</td>
<td>-4</td>
</tr>
<tr>
<td>15</td>
<td>5</td>
<td>9</td>
<td>10</td>
<td>17</td>
<td>14</td>
</tr>
<tr>
<td>-6</td>
<td>-3</td>
<td>-4</td>
<td>-8</td>
<td>-9</td>
<td>-5</td>
</tr>
<tr>
<td>14</td>
<td>6</td>
<td>10</td>
<td>13</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>-6</td>
<td>-4</td>
<td>-2</td>
<td>-9</td>
<td>-5</td>
<td>-7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Borrowing Facts

<table>
<thead>
<tr>
<th>Name</th>
<th>Score</th>
<th>Preferences</th>
<th>Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>467</td>
<td>892</td>
<td>194</td>
<td>678</td>
<td>775</td>
<td>497</td>
</tr>
<tr>
<td>-229</td>
<td>-733</td>
<td>-27</td>
<td>-423</td>
<td>-462</td>
<td>-123</td>
</tr>
<tr>
<td>124</td>
<td>396</td>
<td>521</td>
<td>291</td>
<td>356</td>
<td>778</td>
</tr>
<tr>
<td>-67</td>
<td>-122</td>
<td>-346</td>
<td>-170</td>
<td>-201</td>
<td>-123</td>
</tr>
<tr>
<td>488</td>
<td>187</td>
<td>883</td>
<td>199</td>
<td>796</td>
<td>784</td>
</tr>
<tr>
<td>-263</td>
<td>-39</td>
<td>-629</td>
<td>-21</td>
<td>-682</td>
<td>-753</td>
</tr>
<tr>
<td>789</td>
<td>589</td>
<td>499</td>
<td>876</td>
<td>694</td>
<td>578</td>
</tr>
<tr>
<td>-433</td>
<td>-124</td>
<td>-277</td>
<td>-19</td>
<td>-281</td>
<td>-496</td>
</tr>
<tr>
<td>978</td>
<td>184</td>
<td>286</td>
<td>998</td>
<td>415</td>
<td>482</td>
</tr>
<tr>
<td>-241</td>
<td>-67</td>
<td>-134</td>
<td>-673</td>
<td>-379</td>
<td>-271</td>
</tr>
<tr>
<td>322</td>
<td>497</td>
<td>513</td>
<td>193</td>
<td>892</td>
<td>649</td>
</tr>
<tr>
<td>-198</td>
<td>-327</td>
<td>-369</td>
<td>-72</td>
<td>-161</td>
<td>-327</td>
</tr>
<tr>
<td>387</td>
<td>624</td>
<td>211</td>
<td>396</td>
<td>612</td>
<td>158</td>
</tr>
<tr>
<td>-112</td>
<td>-198</td>
<td>-194</td>
<td>-214</td>
<td>-487</td>
<td>-217</td>
</tr>
<tr>
<td>876</td>
<td>241</td>
<td>801</td>
<td>781</td>
<td>296</td>
<td>827</td>
</tr>
<tr>
<td>-112</td>
<td>-193</td>
<td>-167</td>
<td>-503</td>
<td>-110</td>
<td>-108</td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY

Allyn, T., & Azrin, N. The measurement and reinforcement of adaptive behavior of psychotics. *Journal of the Experimental Analysis of Behavior*, 1965, 8, 357-383.


Moore, J., & Smith W. Knowledge of results in self-teaching spelling. Psychological Reports, 1961, 9, 717-726.


Nickell, M., & Travers, R. Effects of different reinforcers: Comparison across age groups. Psychological Reports, 1964, 13, 739-746.


Rosen, B. The achievement syndrome. In J. Atkinson (Ed.)


Smith, W., & Moore, J. *Size of step and cueing.* *Psychological Reports,* 1962, 10, 287-294. (b)


Thomas, F. *Visualization, experience, and motivation as related to feedback in problem solving.* *American Psychologist,* 1956, 11, 444.

Uhlinger, C., & Stephens, M. *Relation to achievement motivation to academic achievement in students of superior ability.* *Journal of Educational Psychology,* 1960, 51, 259-266.


Wolfensberger, W. Differential rewards as motivating factors in mental deficiency research. *American Journal of Mental Deficiency, 1960, 64, 902-906.*

