THE EFFECTS OF TWO TYPES OF SUCCESS AND FAILURE ON CHILDREN'S ...
MEID, ELIZABETH S.
ProQuest Dissertations and Theses; 1971; ProQuest Dissertations & Theses Global

INFORMATION TO USERS

This dissertation was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

University Microfilms

300 North Zeeb Road
Ann Arbor, Michigan 48106

A Xerox Education Company

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

72-17,144

MEID, Elizabeth S., 1945-
THE EFFECTS OF TWO TYPES OF SUCCESS AND FAILURE ON CHILDREN'S DISCRIMINATION LEARNING AND EVALUATIONS OF PERFORMANCE.

Yale University, Ph.D., 1971
Psychology, experimental

University Microfilms, A XEROX Company, Ann Arbor, Michigan

© Copyright by
ELIZABETH S. MEID
1972

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The Effects of Two Types of Success and Failure on Children's Discrimination Learning and Evaluations of Performance

Elizabeth S. Meid

A Dissertation Presented to the Faculty of
the Graduate School of Yale University
in Candidacy for the Degree of
Doctor of Philosophy
1971

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

PLEASE NOTE:

Some pages may have

indistinct print.

Filmed as received.

University Microfilms, A Xerox Education Company

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Digest

Previous studies of the effects of prior success and failure on children's subsequent learning performance have obtained contradictory findings. The discrepant results may be attributable to differences among these studies in regard to the type of success-failure manipulation employed, the age of children used as subjects, and the possible interaction of these two variables.

The present study attempted to clarify the importance of the effects of two types of prior success-failure manipulations on children's learning and evaluations of performance. The manipulations were either established verbally or defined objectively during two preliminary games. The study examined developmental differences in the effects of the verbal and objective manipulations on children at two age levels, 6 and 10. It investigated directly how the two types of manipulations affected children's feelings of success and failure regarding their past performance on the preliminary games and their anticipated performance on the third game (learning task).

Verbal and objective success-failure manipulations on two experimental games were used with 90 6-year-old and 90 10-year-old children, comprising a 3 X 3 X 2 (Verbal X Objective X Age) factorial design. The verbal success, failure, and control conditions were established respectively by positive comments by the experimenter about the quality of the children's performance on the games, negative comments by the experimenter, and no comments. The objective success, failure, and control conditions were operationally defined respectively by the children's winning a high (80%) or low (20%) percentage of marbles on the games or by not playing the games for marbles. Following the games, the children's feelings of success and failure were assessed directly by means of verbal inquiry and rating scales regarding how well they thought they had done on the games and anticipated doing on the third game (learning task). A three-choice size discrimination learning problem was then presented to the children for 80 trials or until they had reached the criterion of five consecutive correct responses.

No overall effects on learning for the verbal and objective manipulations were found, due in part to a floor effect associated with the greater difficulty of the task when the middle-sized stimulus was correct than when either the large or small stimulus was correct. Among children for whom the large or small stimulus was correct, discrimination learning was better if the adult had made no verbal comments about their performance during the preliminary games than if the adult had verbally evaluated their preliminary performance positively or negatively. Among children who learned the discrimination, boys learned better with negative than with positive or no verbal reinforcement, whereas for girls, the effect was just the opposite. The most striking developmental finding was that in evaluating their past performance, the younger children were sensitive only to the verbal manipulation and did not respond to the objective manipulation, whereas the older children responded to both. Younger girls rated themselves lower than younger boys overall, but there were no differences between the older boys and girls on evaluations of past performance. Additional findings for the older children indicated that high estimates of future performance on the learning task were related to poorer learning, lower IQ, and greater tendency to give socially desirable responses. These findings were discussed in relation to previous research, and areas of future investigation were delineated.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Acknowledgments

The author would like to express her sincerest appreciation to Dr. Susan Harter, chairwoman of the advisory committee, for her constant intellectual and moral support and patient guidance throughout this study. The author is also grateful to Dr. Lita Furby and Dr. Robert Crowder, the other members of the advisory committee, for their thoughtful suggestions and criticisms. Special thanks are due Mrs. Sally Styfco for her invaluable and dedicated services as research assistant, consultant, editor, and typist. The kind co-operation of Mr. Richard Belfonti and Mr. Frederick D'Ambrose, principals of the William Douglas Elementary School and the Stanley T. Williams School in Northford, Connecticut, where the study was conducted, is also gratefully acknowledged.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table of Contents

	Page
Introduction	1
Method	8
Results	19
Discussion	42
References	55
Appendix A	58
Appendix B	62
Appendix C	67

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Introduction

Little systematic research has been conducted investigating developmental differences in children's general reactions to success and failure, in responsiveness to different kinds of success and failure, and the effect of such experiences on children's subsequent learning. Not only would such research aid in understanding the development of success-striving or failure-avoiding orientations to achievement, but would also have important educational implications, e.g., in regard to the structuring of a child's learning experiences and feedback concerning his efforts. While several studies have been addressed to the general problem of the effects of prior success and failure on children's subsequent learning, using both normal and retarded children as subjects, it has been difficult to draw conclusions from the disparate results. Even when the focus has been on normal children only, as in the present study, the findings to date remain contradictory.

The following equivocal findings have been reported in the literature concerning the effects of preliminary success and failure experiences on normal children's subsequent learning performance: Kass and Stevenson (1961) found that a success condition facilitated later performance on a learning task more than either failure or control conditions, with no significant difference found between the latter two conditions. Other studies not utilizing a control condition have found both that prior success facilitated performance (Steigman & Stevenson, 1960) and that it attenuated performance (Stevenson & Zigler, 1958), in comparison to a failure condition. More recently, Butterfield and Zigler (1965) reported that both preliminary success and failure experiences facilitated subsequent performance over a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

control condition. Meid and Zigler (1971) replicated the findings of Butterfield and Zigler that certain conditions of success and failure produced better subsequent learning than a control condition.

It is apparent from this brief summary of studies that findings regarding the effects of preliminary success and failure on normal children's learning are contradictory. However, an examination of the studies reported reveals that they differ on a number of variables. The age or developmental level of the children used as subjects has varied considerably across studies. An important difference has been the manner in which feelings of success and failure have been induced. In certain studies the method has consisted of positive and negative statements by the experimenter regarding the quality of the child's performance, whereas in others, the amount of objective success and failure the child could experience on the preliminary games, e.g., number of marbles won, has been manipulated. Still other studies have used some combination of the two methods. Also, the percentage of reinforcement defining the objective type of success and failure, and thus possibly the magnitude of the intended effect, has varied across experiments.

In addition to these differences, the studies have varied in regard to the nature of the learning task used as a criterion measure. Generally the learning task has been a discrimination learning problem, in which correct responses have been consistently rewarded, although other studies have employed a probability reinforcement procedure. There have also been variations in the response measure employed. Furthermore, some success-failure studies have not included a control group. Those that have included a control group have differed in procedure, in that in some, control subjects

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

were given the learning task with no preliminary activity, and in others, control subjects performed preliminary tasks under conditions having no success-failure implications. These many differences among studies undoubtedly contribute to the contradictory findings and make it difficult to compare them and to draw any consistent conclusions.

An important variable in understanding the effects of prior success and failure on children's subsequent learning is the nature of the experimental success-failure manipulation. Previous investigators have not considered or identified different types of success-failure manipulations nor compared their relative effectiveness. Two major types of qualitatively different success-failure experiences can be distinguished. Feelings of success or failure can be established on the one hand by another person's evaluation of one's performance or, on the other hand, by one's own evaluation of his performance based on objective indicators of success and failure. A major purpose of the present study was to investigate the relative influence of each type of success-failure experience on children by utilizing two different methods of inducing success and failure. One consisted of positive or negative statements by the experimenter about the quality of the child's performance on two games prior to the learning task. The other consisted of varying the amount of objective success (number of marbles won) experienced by the child on the games. While both types of manipulations have been used in previous studies, no study has systematically examined their effects separately, cumulatively, and in opposition to each other in the same study. Consequently, in the present study some subjects experienced success determined solely by the verbal method, some solely by the objective method, and some by a combination of both methods. Other subjects

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

experienced comparable verbal and objective failure conditions. Groups were included in which verbal success was paired with objective failure and objective success with verbal failure, to assess in the same subjects the relative influence of competing methods of success-failure manipulations. There was also a "pure" control condition with no success-failure manipulations.

A second major purpose of the present study was to investigate whether type of success-failure manipulation might interact with age to produce developmental differences in children's reactions to different methods of inducing feelings of success and failure. Consequently, unlike previous studies which have not included the developmental dimension, the present study included children at two different age levels, 6 and 10. The demonstration of developmental differences in the motivational and personality structure of children supports the hypothesis that a child of a certain age may be particularly sensitive to either a predominantly verbal or objective type of success and failure. In the area of social reinforcement, for example, a developmentally changing verbal reinforcer hierarchy has been postulated, in which praise is a more effective reinforcer for young children, but correctness the more effective dimension for older children (Harter, 1967; Stevenson, 1961, 1965). Other research on children's locus of control indicates that with increasing age, children tend to shift from external to internal locus of control; i.e., older children tend more to perceive the outcome of events as being internally, rather than externally, controlled than do younger children (Cromwell, 1963). Findings with the Children's Social Desirability scale also show a shift from a greater to a lesser tendency to give socially desirable responses with increasing age,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

suggesting a decrease in need for social approval from an adult (Crandall, Crandall, & Katovsky, 1965).

These findings in the areas of social reinforcement, personality development, and motivation support the tentative hypothesis of a developmental shift in effectiveness of verbal and objective types of success-failure manipulations. In particular, on the basis of this research, it seems likely that the younger the child, the more influenced he is apt to be by another person's evaluation of his performance. Objective assessment of one's own performance would seem to be a capacity which develops at a later stage than reaction to the evaluation of others. Thus one would predict that a verbal manipulation would be more effective than an objective manipulation with younger children, whereas an objective manipulation would have a greater effect on older than on younger children.

A third major purpose of the present study was to overcome a major limitation of previous studies of the effects of prior success and failure on children's subsequent learning, namely, the failure to assess directly the effects of the manipulations by measuring the child's feelings of success or failure on the experimental tasks and also his expectancy of succeeding or failing on the learning task, as related to the manipulations. In previous studies, the effects of the manipulations have been only inferred from subsequent performance on the learning task or from incidental behavioral observations during the manipulations. Consequently, in the present study, the child's feelings about his past and future performance were investigated by verbal inquiry and rating scales following the experimental games. Thus it was possible to assess directly the effects of the manipulations by asking the child about his feelings and also by examining

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the relationship between the child's stated feelings of success and failure and his learning performance. In addition, the effects of the manipulations could be inferred, as in previous studies, from learning performance.

A minor consideration of the present study was to examine the older children's scores on the Children's Social Desirability scale, measuring need for adult approval by the tendency to give socially desirable responses (Crandall et al., 1965), in relation to their stated feelings about past performance on the experimental games and future performance on the learning task. The purpose of this examination was to investigate whether the children's evaluations of their performance reflected a more global personality trait such as the tendency to give socially desirable responses. Thus, the Children's Social Desirability scale was administered to the older children in this study. The younger children were excluded from this measure, since it is questionable whether children below the third grade, which is the lowest age level used in previous studies, comprehend the scale.

Thus the basic design of the present study was a 3 X 3 X 2 factorial which included verbal success-failure manipulations, objective success-failure manipulations, and age. In some instances, sex differences were another variable (see Table 1). Three verbal manipulation conditions and three objective manipulation conditions were established during two preliminary games. In the verbal manipulation, success, failure, and control conditions were established respectively by positive statements by the experimenter, negative statements by the experimenter, and by the experimenter's remaining silent (subsequently called the Positive Comment, Negative Comment, and No Comment conditions). In the objective manipulation, success,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

failure, and control conditions were operationally defined respectively by high and low percentages of marbles won on the preliminary games and by the subject's playing with the game materials for no marbles (subsequently called the 80% Reinforcement, 20% Reinforcement, and No Marble conditions). Subjects at each of two age levels (6 and 10) were included in each experimental group.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Method

Subjects

Each of the 18 groups in the 3 X 3 X 2 design contained 10 subjects. The mean chronological age (CA) of the younger subjects was 6 years, 8 months (standard deviation 3.6 months) and of the older subjects was 10 years, 9 months (standard deviation 4.7 months). CA means and standard deviations were matched across groups at each age level. Subjects were obtained from nine classrooms in two elementary schools in Northford, Connecticut. Only white, middle-class children from regular first- and fifth-grade classrooms were included in the study. Subjects were assigned to conditions in such a way as to control for any possible effects of sex, classroom, or teacher. Derived IQ scores on the Otis-Lennon Mental Ability Test were available for the fifth-graders, and percentile scores on the Metropolitan Reading Readiness Test were available for most of the first-graders. The mean IQ of the fifth-graders was 112.4, and the mean percentile score of the first-graders was 61.1. These scores were not confounded with any experimental conditions.

Experimenters

Two young women having considerable experience in the testing of children each tested half of the subjects in each of the 18 groups. The first was an advanced graduate student in the Yale Department of Psychology and author of this study (Experimenter 1). The other was an assistant in research employed by the Yale Department of Psychology (Experimenter 2). The experimenters practiced to be as similar as possible in their procedure, including tone of voice, facial expression, and interaction with subjects prior to completion of the criterion task.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

General Procedure

Subjects in all but the No Comment-No Marble group played two preliminary games, first Pick-A-Card and then Which-School, after which they received the criterion task, a three-choice size discrimination learning problem. Directly following the preliminary games, the experimenter questioned the subject concerning his strategy and performance on the games, and his estimate of future performance on the third game; she also had the subject mark scales indicating how well he thought he did on the games and how well he thought he would do on the third game. These measures were included to assess the effectiveness of the success-failure manipulations. Subjects in the No Comment-No Marble group were given an opportunity to play with the game materials for the average amount of time (7 minutes) that other subjects took playing the games, and were questioned only regarding the general nature of what they had done with the materials and how well they thought they would do on the criterion task. The reason that these subjects did not merely play the games in the absence of verbal or objective reinforcement was that pretesting indicated that such a procedure still involved possible success-failure implications. Throughout the procedure up to the learning task, instructions to subjects in the No Marble groups differed from those in the 80% Reinforcement and 20% Reinforcement groups in that no mention was made of playing the preliminary games for marbles or of marble boards.

Each subject was taken individually from his classroom to a testing room. When the subject entered the room, the experimenter gestured to a variety of attractive toys or objects appropriate for the subject's age and sex and said, "See, I have some toys and things here. I'm going to give

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

you a chance to win one of these today. Which one do you like best? Which one would you most like to have?" After the subject made his choice, the experimenter said to subjects in the 80% Reinforcement and 20% Reinforcement conditions only:

> Fine. Now we're going to play some games. By playing the games well, you can win marbles. If you win enough marbles, then you can have this. Sometimes you'll win and sometimes you won't. Each time you win, you'll get a marble to put in a marble board. If you win every time, you'll be able to fill all the holes in the marble board.

To subjects in the Positive Comment-No Marble and Negative Comment-No Marble groups, the experimenter said, "By playing the games well, you can win the toy." Subjects in the No Comment-No Marble group were told:

> Later I'm going to give you a chance to win this toy. We'll play a game where you can win marbles, and if you win enough, then you can have the toy. But first, I have to finish some things here. While I'm doing that, you can play with these things here.

Then, for all subjects but those in the No Comment-No Marble group, the experimenter introduced the first preliminary game. After the first game, the experimenter presented the materials and gave the instructions for the second game. Following this game, the measures designed to assess the effectiveness of the success-failure manipulations were given.

Up to this point in the procedure, subjects in the No Comment-No Marble group had had the opportunity to play quietly with the game materials while the experimenter busied herself with papers. After 7 minutes, the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

experimenter said, "Okay, we're almost ready. Would you please help me pick these things up? We'll put the pictures in the boxes and pick up the cards." After the materials were removed, the experimenter conducted the inquiry and introduced the criterion task.

The learning task was then presented to all subjects, and the subject played it until he reached a criterion of five consecutive correct responses or 80 presentations of the stimuli. Following the learning task, subjects were questioned about their strategy on the game, to see if they could verbalize the solution to the learning problem. At this point, subjects in the five groups having failure aspects (Negative Comment-80% Reinforcement, 20% Reinforcement, No Marble; Positive Comment-20% Reinforcement; No Comment-20% Reinforcement) replayed part of the preliminary games under high success conditions and were praised extensively for the "improvement" in their performance, to eliminate any possible negative effects of the manipulations. Then the subject counted the marbles he had won and was told that he had more than enough to win the prize, which would be given to him right before he left school that day. Each subject was then praised liberally and returned to his classroom. In addition, each subject was asked not to talk about the exact nature of the games to any of the other children until after everyone had played.

Approximately 10 days after all subjects had been tested, two different female research assistants administered the Children's Social Desirability scale to the fifth-graders in their classroom groups.

Experimental Games

The two experimental games used in this study were adapted from those used by Stevenson and Zigler (1958). Both games were quite simple, and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

performance depended primarily upon compliance with the experimenter's instructions.

Pick-A-Card. In this game, the subject's task was to hand to the experimenter 20 1-inch (25.4mm) X 3-inch (76.2mm) cards each having a red or a blue spot, one at a time. As the cards were being spread before the subject, he was told:

> Okay. Now here's the first game. It's called Pick-A-Card. Here are the cards. Some have blue spots and some have red spots. Your job is to pick up cards whose spots are the same as the colors I have written down here on this piece of paper. Sometimes a blue one is right and sometimes a red one is right. I'll spread the cards all over the table, like this. When it's time for you to begin, hand the cards to me one at a time.

At this point in the instructions, subjects in the No Marble groups were told, "All right, now you may begin." To the subjects in the 80% Reinforcement and 20% Reinforcement conditions, the experimenter continued:

> Every time you give me a right-colored card, I'll give you a marble. If you have enough marbles after we have played three games, you can have this (indicating prize). You can keep the marbles for this game in this marble board, and we'll count all the marbles after we have played three games to see if you have won enough. All right, now you may begin handing the cards to me one at a time. When you give me a right card, I'll give you a marble to put in your board. Now let's play, and you try to get as many marbles as you can.

Which-School. In this game, the subject's task was to place pictures of children, mounted on three different shapes of construction paper of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

three different colors, into either of two open boxes, each painted a different color both inside and outside. The boxes were placed about 8 inches (203.2mm) apart and 6 inches (152.4mm) in front of the subject with their open sides facing him. The experimenter said, "Okay. Now here's the next game." To subjects in the 80% Reinforcement and 20% Reinforcement conditions, the experimenter added, "You get a new marble board for this game" and gave it to him. Then for all subjects, the experimenter continued:

> It's called Which-School. Here are the two schools. One is red and one is blue. Here are some pictures of children. Each of these children goes to one of these schools. Your job is to figure out which school each child goes to. Take the pictures one at a time and put them in the school you think each child goes to.

At this point, subjects in the 80% Reinforcement and 20% Reinforcement conditions were told, "When you put the picture in the right school, I'll give you a marble to put in your board." The experimenter finished the instructions by saying, "Here are the pictures. You put them in the schools one at a time."

Manipulations

The success, failure, and control conditions in the verbal manipulation were established during the preliminary games in the following way: The success (Positive Comment) condition was defined by four positive statements made to the subject by the experimenter at predetermined points during each of the preliminary games. In the 80% Reinforcement and 20% Reinforcement conditions, these statements always occurred on a response for which the subject was given a marble. The positive statements (with

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

parentheses indicating additions made on the second game) were: "Yes," "That's good," "You play this game very well (too)," "You're (still) playing really well. Not many of the other kids got as many right as you have (on this game too)." The failure (Negative Comment) condition was likewise defined by four negative statements made by the experimenter at predetermined points during both preliminary games. In the marble (80% Reinforcement and 20% Reinforcement) conditions, these statements always occurred following a response for which the subject did not receive a marble. The negative statements were: "No," "That's not good," "You don't play this game very well (either)," "You're (still) playing pretty poorly. Most of the other kids got every single one of them (these) right (too)." In the control (No Comment) condition, the experimenter made no statements to the subject during the games.

The success, failure, and control conditions in the objective manipulation were established during the preliminary games by the percentage of marbles the subject won or by playing the games for no marbles. Subjects in the success (80% Reinforcement) condition won 16 of a possible 20 marbles on each game. Subjects in the failure (20% Reinforcement) condition won 4 of a possible 20 marbles on each game. A marble board, made of wood 6 X 9 X 3/4 inches (152.4 X 228.6 X 19.1mm) painted gray and having 20 holes arranged 4 X 5, was used to hold the marbles won during each game. At the end of each game, the marble board was removed but remained in the subject's sight throughout the rest of the procedure. Thus at the end of the games, subjects in the 80% Reinforcement condition had won 32 marbles which filled 16 of the 20 holes in each board. Subjects in the 20% Reinforcement condition had won 8 marbles which filled only 4 of 20 holes in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

each board. Subjects in the control (No Marble) condition either played the preliminary games (Positive Comment-No Marble and Negative Comment-No Marble groups) or played with the game materials (No Comment-No Marble group) for no marbles.

On each game in all success and failure conditions, the various reinforcements were administered once in each block of five trials. In the marble conditions, trials for positive and negative reinforcement were selected so that there were at least two nonreinforced trials between the absence of a marble (80% Reinforcement condition) or the presence of a marble (20% Reinforcement condition). In the verbal (Positive Comment and Negative Comment) conditions, positive statements could not follow or precede directly the absence of a marble, and negative statements could not follow or precede directly the presence of a marble.

Questions and Bar Scales

At the end of the second preliminary game, the experimenter questioned each subject (except those in the No Comment-No Marble group) about the games in an attempt to determine the extent to which the subject had formulated a strategy on them and the effectiveness of the success-failure manipulations. The experimenter asked about what the subject was doing during the games, how well he thought he did, if he thought he won enough marbles or played well enough to get the prize, and how well he thought he would do on the third game. In addition, the subject marked each of two 8-bar scales to indicate how well he thought he had done on the games and how well he thought he would do on the third game. Subjects in the No Comment-No Marble group were asked about what they were doing with the materials and marked a bar scale to indicate how well they thought they

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

would do on the criterion game. The exact questions the experimenter asked subjects in the various conditions and instructions regarding the bar scales are included in Appendix A.

Following the learning task, the experimenter asked the subject what he thought about the game, what he thought he was supposed to do in it, and how he decided which block to pick each time, to see if the subject could verbalize the solution to the learning problem. The exact questions are included in Appendix A.

Learning Task

Although the learning task was presented to the subjects as just a game in which they could win marbles, it differed from the preliminary games in that securing marbles depended upon learning a correct response. The criterion measure was a three-choice size discrimination learning problem. The subject's task was to find marbles by picking up one of three different-sized blocks each time they were presented to him.

The stimuli were square blocks cut from 1/2-inch (12.7mm) plyboard and painted flat black. Pretesting indicated that different sizes were necessary at the two age levels to equate the difficulty of the task at both ages. For the younger subjects, the blocks were 3, 4 1/2, and 6 inches (76.2, 114.3, and 152.4mm) square; for the older subjects, blocks, 3, 3 1/4, and 3 1/2 inches (76.2, 82.6, and 88.9mm) square were used. The apparatus was a modification of the Wisconsin General Testing Apparatus. It consisted of a 14 X 16-inch (355.6 X 406.4mm) wood tray having three wells 7/8-inch (22.4mm) deep and 4 3/4 inches (120.7mm) apart. The stimuli covered the wells, one of which on each trial contained a marble. The tray was pushed towards the subject from beneath a 6 X 18-inch (152.4 X 457.2mm) curtain.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The curtain prevented the subject from seeing either the stimuli or the experimenter's hand movements when the tray was pulled toward the experimenter. The subject was allowed to pick up and look under only one block on each presentation. A noncorrection procedure was used, and a trial was defined by lifting a block. The arrangement of the blocks was changed after each presentation according to a predetermined schedule such that within each block of six trials, each possible stimulus arrangement appeared once. No stimulus appeared more than twice in a row in the same position, and the arrangement was further modified so that any position response strategy was not reinforced. The stimulus to be rewarded was randomly determined for each subject, with the restriction that each stimulus be correct for approximately one-third of the subjects in every group. The correct stimulus for each subject was 100% reinforced.

The criterion task was presented to all subjects in the following manner. Additions for subjects in the 80% Reinforcement and 20% Reinforcement (marble) conditions are indicated by parentheses.

> Okay, now we're going to play the last game. It's called Find-the-Marble. By playing well, you can win lots of marbles. If you win enough marbles (remember), you can have this prize. Now let me show you how to play this game. I am going to hide marbles under these blocks. You try to find them. Every time you find a marble, you can put it in this marble box. At the end of the game, if you have found enough marbles (together with the ones in these boards), you can trade them to me for the toy that you want. Every time I push the blocks out, you can look underneath one of them.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(For subjects in the No Comment-No Marble group, the only difference from these instructions was to refer to "the game" rather than "the last game.")

Marbles were placed in a transparent container as they were found. Nothing was said to any subject during the criterion task except to repeat instructions about picking up only one stimulus or that it was "no fair" to try to watch under which block the experimenter hid the marble.

Children's Social Desirability Scale

In order to assess the differing tendencies among the older subjects to give socially desirable responses, the direct-question form of the Children's Social Desirability scale (Crandall et al., 1965) was employed. Test-retest reliability has been found to be high, .90. The Children's Social Desirability scale presents a child with 47 "yes-no" items stated so that a subject cannot answer them in a socially desirable manner without dissembling. Some of the items ask a child if he always behaves in some particular fashion which is prescribed by middle-class American mores or always holds such attitudes and beliefs, e.g., "Are you always polite to older people?" Other items ask if he never does, or thinks of doing, those things which are disapproved of in his culture. "Have you ever felt like saying unkind things to a person" is representative of this kind of item. Still other items pose an unacceptable behavior or attitude and ask whether he sometimes acts or thinks in this fashion, e.g., "Have you sometimes felt like throwing or breaking things?" The social desirability score is determined by the number of items on which the subject maintains that he has such an undeviating socially desirable attitude or behavior (13 items requiring a "yes" and 34 items requiring a "no" response in order to obtain the maximum social desirability score).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Results

Analyses were performed to investigate the effects of the experimental manipulations on both learning performance on the discrimination task and on subjects' feelings of success and failure as reflected in estimates of past performance on the experimental games and future performance on the learning task. Further analyses were performed to examine the relationship between subjects' estimates of past and future performance and their performance on the learning task. In addition, some relationships among estimates of past and future performance of the older subjects, their IQ and Children's Social Desirability scale scores, and their learning performance were examined.

Effects of Prior Success-Failure Manipulations on Discrimination Learning Performance

Two measures of learning on the discrimination task were employed, total number of correct responses and number of trials to criterion. For subjects who reached criterion, the score for number of correct responses included the number of trials beyond the criterion block up to 80 (at which point the task was terminated for all subjects), based on the assumption that all responses beyond criterion would be correct. The score for trials to criterion included the five criterion trials. Number of correct responses and trials to criterion were found to be highly negatively correlated for learners ($r = -.97$, $p < .001$). Only results based on number of correct responses will be presented, thus enabling all subjects, including those who failed to reach criterion (48 of 180 subjects), to be included in the overall analyses. (Summary F tables and tables of means for analyses based on trials to criterion are presented in Appendix B.)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The expectation of no learning differences associated with age, derived from pretesting with stimuli of different dimensions for the two age levels, was confirmed in preliminary overall analyses.[1] Among the younger children, 64 learned and 26 did not; of the older children, 68 learned and 22 did not. These differences were not significant ($\chi^2 = .26$), further confirming that the task was of intermediate difficulty at both ages, avoiding floor and ceiling effects, and thus appropriate for revealing differences in learning.

Learning performance of all subjects (learners and nonlearners combined). Two 3 X 3 X 2 X 2 (Objective X Verbal X Age X Sex, and Objective X Verbal X Age X Experimenter) analyses of variance performed on the number of correct responses revealed no significant effects.[2] Means involved in the Objective X Verbal X Age X Sex analysis are presented in Table 1.

Inspection of the data indicated that the absence of overall effects may have been related to the greater difficulty of the discrimination learning task when the middle-sized stimulus was correct than when the large or small stimuli were correct. Fewer correct responses were made to the middle-sized stimulus ($\bar{X} = 42.0$) than to the large ($\bar{X} = 57.4$, $p < .01$) or small ($\bar{X} = 61.6$, $p < .01$) stimuli.[3] This was revealed in a 3 X 2 X 3 (Objective X Age X Stimulus Size) analysis of variance by a

---

1. The only exception was a slight age effect revealed in an analysis of trials to criterion for learners only. See Appendix B for the summary table of this analysis.

2. An overall 5-way analysis of variance including both the Sex and Experimenter variables was not performed because of the small Ns (e.g., 2) in certain cells.

3. All multiple mean comparison tests employed the Newman-Keuls method, unless otherwise indicated.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 1

Mean Number of Correct Responses for All Subjects

(and for Only Subjects for Whom Large or Small Stimuli Were Correct)

in Each Group at Each Age by Sex

Objective Conditions	Verbal Conditions					
	Positive Comment		Negative Comment		No Comment	
	Male	Female	Male	Female	Male	Female
80% Reinforcement						
Age 6	57.8 (56.8)	35.5 (47.5)	60.5 (66.0)	41.8 (48.0)	61.2 (61.2)	47.3 (65.0)
Age 10	52.4 (60.3)	47.0 (49.5)	53.6 (60.5)	44.8 (50.0)	54.4 (66.7)	51.0 (48.8)
20% Reinforcement						
Age 6	48.8 (59.0)	40.3 (63.0)	55.8 (62.2)	59.3 (69.7)	53.2 (75.7)	54.3 (72.0)
Age 10	52.2 (53.0)	56.4 (49.0)	58.8 (55.8)	62.2 (62.0)	48.2 (67.5)	59.8 (58.3)
No Marble						
Age 6	49.3 (53.0)	55.0 (62.0)	63.8 (74.3)	32.3 (33.3)	38.2 (53.0)	66.0 (66.0)
Age 10	45.6 (50.0)	63.6 (73.0)	63.4 (53.3)	50.8 (50.8)	63.6 (74.7)	58.4 (74.0)
$(\bar{X})$	(56.5)		(57.8)		(64.1)	

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

statistically significant main effect for Stimulus Size ($F_{2,162}$ = 20.64, $p < .01$). Furthermore, the discrimination was more difficult to learn with the middle-sized stimulus than with the other stimuli for younger children than for older, as reflected in the Age X Stimulus Size interaction ($F_{2,162}$ = 4.22, $p < .05$) in this analysis. The means for this interaction are presented in Table 2. At Age 6, the middle-sized stimulus received fewer correct responses than the small ($p < .01$) or large ($p < .01$) stimuli, but at Age 10 it received fewer correct responses than the small stimulus only ($p < .05$). Thus for the younger children the middle-sized stimulus was harder to learn than the other two, but for the older children it was harder than the small stimulus only. Very similar results were obtained from a 3 X 2 X 3 (Verbal X Age X Stimulus Size) analysis of variance, which revealed a Stimulus Size main effect ($F_{2,162}$ = 20.01, $p < .01$) and an Age X Stimulus Size interaction ($F_{2,162}$ = 4.06, $p < .05$). Again, the Stimulus Size main effect was due to the greater difficulty of the task with the middle-sized stimulus than with the large or small stimuli, and the Age X Stimulus Size interaction reflected the greater difficulty of the task with the middle-sized stimulus for the younger children than for the older children.

Inspection of the number of subjects who learned the discrimination showed that 28 out of 61 learned with the middle-sized stimulus and 104 out of 119 learned with the large or small stimuli. To analyze further the differences in learning associated with stimulus size, a chi-square test was performed on the number of learners and nonlearners among all subjects assigned the middle-sized stimulus or the small and large stimuli combined as correct. The results indicated that the discrimination was more difficult

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 2

Mean Number of Correct Responses for All Subjects at Each Age with Each Stimulus Size as Correct

Age	Stimulus Size		
	Small	Middle-sized	Large
6	60.4	35.2	59.7
10	62.6	48.6	55.0

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to learn with the middle-sized stimulus (χ^2 = 40.08, p < .001). Of subjects for whom the middle-sized stimulus was correct, 8 learned and 22 did not at Age 6, and 20 learned and 11 did not at Age 10 (χ^2 = 7.34, p < .01); thus the discrimination was more difficult to learn at Age 6 with the middle-sized stimulus as correct than it was at Age 10.

As a result of the generally greater difficulty of the task with the middle-sized stimulus at both ages, those cells which contained a disproportionately greater number of subjects assigned the middle-sized stimulus were more likely to reflect poorer learning. Inspection of the data showed that all Sex X Experimenter interactions would be artifacts of this unequal distribution in that a greater number of middle-sized stimuli were present in certain of the cells involved in the Sex X Experimenter breakdowns. For this reason, in addition to the small Ns in some of the cells, analyses involving both the Sex and Experimenter variables in the same analysis are not presented.

Learning performance of subjects for whom the large and small stimuli were correct. Because the learning task was more difficult with the middle-sized stimulus at both age levels, as revealed by the analyses reported above, a floor effect associated with the middle-sized stimulus may have obscured the effects of the manipulations on learning performance. Consequently, a 3 X 3 X 2 X 2 (Objective X Verbal X Age X Sex) analysis of variance of number of correct responses was performed omitting those subjects who had received the middle-sized stimulus as correct. Means for this analysis are presented in Table 1. A Verbal main effect was revealed ($F_{2,83}$ = 3.45, p < .05), indicating that subjects in the No Comment condition made more correct responses ($\bar{X}$ = 64.1) than subjects in the Positive

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Comment ($\bar{X}$ = 56.5) or Negative Comment ($\bar{X}$ = 57.8) conditions. No other significant effects were revealed.

Learning performance of subjects who learned the discrimination. To examine further the effects of the experimental manipulations on learning, analyses were performed on scores obtained only by all subjects who learned the discrimination.[1] A 3 X 3 X 2 X 2 (Objective X Verbal X Age X Sex) analysis of variance of number of correct responses revealed a significant interaction between the Verbal conditions and Sex ($F_{2,96}$ = 3.36, p < .05). The means involved in the interaction are presented in Table 3. The Verbal X Sex interaction reflected the fact that while boys learned better in the Negative Comment than in the Positive Comment and No Comment conditions, girls learned worse in the Negative Comment than in the other two conditions. A 3 X 3 X 2 X 2 (Objective X Verbal X Age X Experimenter) analysis of variance of number correct for all learners revealed a significant Age X Experimenter interaction ($F_{1,96}$ = 5.70, p < .05). The means involved in the interaction are presented in Table 4. The interaction reflected the fact that the younger group made more correct responses with Experimenter 2 than with Experimenter 1, but that there were no differences in learning associated with experimenter in the older group.

Effects of Success-Failure Manipulations on Estimates of Past and Future Performance

To examine the effects of the manipulations on subjects' feelings of success and failure, analyses were performed on the older subjects' self-

1. The results of the analyses for learners having only the large and small stimuli as correct were very similar to those for all learners, and because cell Ns in the former analyses were quite small, only results obtained with all learners will be presented.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 3

Mean Number of Correct Responses for Learners Only in Each Verbal Condition by Sex

Sex	Positive Comment	Negative Comment	No Comment
Male	59.6	69.2	62.0
Female	63.5	59.4	64.2

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 4

Mean Number of Correct Responses for Learners Only at Each Age for Each Experimenter

Age	Experimenter 1	Experimenter 2
6	58.5	65.5
10	66.1	62.8

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ratings on the 8-point bar scales of their estimates of past performance on the two experimental games and of their future performance on the third game. The younger subjects' ratings on these scales could not be used, for in conducting the study it was apparent that the younger children did not understand the scales. However, it was possible for the two experimenters together to rate on a 3-point estimate of past performance scale the younger subjects' responses to the first question of the post-experimental game inquiry: "How well do you think you did on the games?" Responses of the younger subjects pertaining to estimates of future performance were not scorable and there is no data on this measure for them. On all scales, a higher rating indicates an estimate of better performance. Many statistically significant results were obtained, and in order to facilitate the presentation of the findings, the F values are not presented in the text but appear in tables in Appendix C. Unless otherwise indicated, all results of analyses of variance in this section were significant at the $p \leq .05$ level or better.

Estimate of past performance: younger subjects. Since data on estimate of past performance were lacking in the No Comment-No Marble group, overlapping 3 X 2 and 2 X 3 analyses of variance omitting this group were performed on the estimate of past performance scales for each age group separately, since the scales were not comparable. Preliminary inspection of the data suggested a possible sex effect associated with the 3-point estimate of past performance scale for the younger subjects, and thus sex was included as a variable in analyses of this scale. The mean estimates of past performance of younger subjects in each of the Verbal X Objective groups except the No Comment-No Marble group are presented in Table 5.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 5

Mean Scores of Younger Subjects in Each Group on

3-Point Estimate of Past Performance Scale

Objective Conditions	Verbal Conditions		
	Positive Comment	Negative Comment	No Comment
80% Reinforcement	2.8	1.6	3.0
20% Reinforcement	2.8	1.7	2.4
No Marble	3.0	1.7	---

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A 3 X 2 X 2 (Verbal X Objective X Sex) analysis of variance omitting the No Marble condition performed on these estimates revealed that the younger subjects estimated their past performance lower in the Negative Comment ($\bar{X}$ = 1.61) than in the No Comment ($\bar{X}$ = 2.68, $p < .01$) or Positive Comment ($\bar{X}$ = 2.80, $p < .01$) conditions, as indicated by a significant Verbal main effect. The No Comment and Positive Comment conditions did not differ. In addition, girls rated themselves generally lower ($\bar{X}$ = 2.24) than boys ($\bar{X}$ = 2.46), as revealed by a significant main effect for Sex. No other effects were significant. A 2 X 3 X 2 (Verbal X Objective X Sex) analysis of variance omitting the No Comment condition performed on estimates of past performance of the younger subjects showed again that they estimated their performance worse in the Negative Comment ($\bar{X}$ = 1.64) than in the Positive Comment ($\bar{X}$ = 2.87) condition, as indicated by a significant Verbal main effect. No other effects were significant.

Estimate of past performance: older subjects. For the older subjects, preliminary analyses of estimates of past performance including sex as a variable revealed no significant differences associated with it, and for the purpose of clarity this variable was omitted from subsequent analyses.[1] The mean scores of older subjects in each of the Verbal X Objective groups except the No Comment-No Marble group, for which there was no score, on the 8-point estimate of past performance scale are presented in Table 6. A 3 X 2 (Verbal X Objective) analysis of variance omitting the No Marble condition performed on these estimates revealed that the older subjects rated themselves lower in the Negative Comment ($\bar{X}$ = 3.30) than in the No Comment

1. Summary tables of the results of the analyses including sex as a variable are included in Appendix C.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 6

Mean Scores of Older Subjects in Each Group on

8-Point Estimate of Past Performance

(and Estimate of Future Performance) Scales

Objective Conditions	Verbal Conditions			
	Positive Comment	Negative Comment	No Comment	$\bar{X}$
80% Reinforcement	6.2 (5.5)	4.9 (5.6)	5.9 (6.1)	5.67 (5.73)
20% Reinforcement	5.1 (5.9)	1.7 (2.9)	4.3 (5.1)	3.70 (4.63)
No Marble	6.0 (5.7)	3.0 (4.7)	— (6.2)	---- (5.53)
$\bar{X}$	5.77 (5.70)	3.20 (4.40)	--- (5.80)	

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

($\bar{X}$ = 5.10, $p < .01$) or Positive Comment ($\bar{X}$ = 5.65, $p < .01$) conditions, and lower also in the 20% Reinforcement ($\bar{X}$ = 3.70) than in the 80% Reinforcement ($\bar{X}$ = 5.67) condition, as indicated by statistically significant Verbal and Objective main effects. The differences between estimates in the Negative Comment compared to the Positive Comment and No Comment conditions were greater in the 20% Reinforcement (both differences significant at $p < .01$) than in the 80% Reinforcement condition (Negative Comment-80% Reinforcement group lower than Positive Comment-80% Reinforcement group at $p < .05$), as revealed by a statistically significant Verbal X Objective interaction. Likewise, the difference between estimates in the 20% Reinforcement and 80% Reinforcement conditions was greater in the Negative Comment ($p < .01$) and No Comment ($p < .01$) conditions than in the Positive Comment condition.

In a 2 X 3 (Verbal X Objective) analysis of variance omitting the No Comment condition, past performance estimates of the older subjects were lower in the Negative Comment ($\bar{X}$ = 3.20) than in the Positive Comment ($\bar{X}$ = 5.77) condition, lower in the 20% Reinforcement ($\bar{X}$ = 3.40) than in the No Marble ($\bar{X}$ = 4.50) condition ($p < .01$), and lower in the No Marble than in the 80% Reinforcement ($\bar{X}$ = 5.55, $p < .01$) condition. The difference between estimates in the Positive Comment and Negative Comment conditions was greater in the 20% Reinforcement ($p < .01$) and No Marble ($p < .01$) conditions than in the Positive Comment condition ($p < .05$), as reflected by a statistically significant Verbal X Objective interaction. Likewise, the differences in estimates among the 80% Reinforcement, 20% Reinforcement, and No Marble conditions were greater in the Negative Comment (all comparisons significantly different at $p < .01$) than in the Positive Comment

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

condition (Positive Comment-20% Reinforcement group lower than Positive Comment-80% Reinforcement group at $p < .05$).

Estimate of future performance: older subjects. Preliminary analyses of the older subjects' estimates of future performance on the learning task which included sex as a variable revealed no statistically significant differences, and for the purpose of clarity, this variable was omitted from subsequent analyses.[1] The mean scores of older subjects in each of the Verbal X Objective groups on the 8-point estimate of future performance scale are presented in Table 6. A 3 X 3 (Verbal X Objective) analysis of variance of these estimates revealed that the older subjects in the Negative Comment condition expected to do more poorly on the learning task ($\bar{X} = 4.40$) than subjects in the Positive Comment ($\bar{X} = 5.70$, $p < .01$) or No Comment ($\bar{X} = 5.80$, $p < .01$) conditions, as indicated by a significant Verbal main effect. Furthermore, subjects in the 20% Reinforcement condition expected to do worse ($\bar{X} = 4.63$) than subjects in the 80% Reinforcement ($\bar{X} = 5.73$, $p < .01$) or No Marble ($\bar{X} = 5.53$, $p < .01$) conditions, as revealed by a significant Objective main effect. A significant Verbal X Objective interaction was also revealed, reflecting the fact that the differences between the Negative Comment and Positive and No Comment conditions were most pronounced in the 20% Reinforcement condition, less in the No Marble condition, and negligible in the 80% Reinforcement condition. Likewise, the differences between the 20% Reinforcement and the 80% Reinforcement and No Marble conditions were greatest in the Negative Comment condition, moderate in the No Comment condition, and very slight in the

---

1. Summary tables of the results of the analyses including sex as a variable are included in Appendix C.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Positive Comment condition. The only statistically significant differences were between the Negative Comment-20% Reinforcement group and all other groups (all at $p < .01$). No other significant effects were found.

Comparison of estimate of past and future performance scales for older subjects. Estimates of past performance on the experimental games and of future performance on the learning task of the older subjects were highly correlated ($r = .74$, $p < .01$). However, the experimental manipulations influenced estimates of past performance more than estimates of future performance, as was revealed in part by a significant main effect for Scale in a 3 X 2 (Verbal X Objective) repeated measures analysis of variance omitting the No Marble condition. Estimates of past performance were generally lower ($\bar{X} = 4.68$) than estimates of future performance ($\bar{X} = 5.18$). However, this difference occurred only in the Negative Comment ($p < .01$), No Comment ($p < .05$), and 20% Reinforcement ($p < .01$) conditions, as indicated by significant Verbal X Scale and Objective X Scale interactions. Estimates of past and future performance in the Positive Comment and 80% Reinforcement conditions did not differ. The means for the Verbal X Scale interaction are presented in Table 7, and those for the Objective X Scale interaction are included in Table 6. No other significant effects were found.

Similar results obtained from a 2 X 3 (Verbal X Objective) repeated measures analysis of variance omitting the No Comment condition confirmed that the experimental manipulations influenced estimates of past performance more than estimates of future performance, in a downward direction, but in certain groups more than in others. Thus a significant Scale main effect indicated that again, estimates of past performance were generally lower ($\bar{X} = 4.48$) than estimates of future performance ($\bar{X} = 5.05$), but significantly

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 7

Mean Scores of Older Subjects in Each Verbal Condition

Omitting No Marble Objective Condition

on 8-Point Estimate of Past Performance

(and Estimate of Future Performance) Scales

Verbal Conditions		
Positive Comment	Negative Comment	No Comment
5.65	3.30	5.10
(5.70)	(4.25)	(5.60)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

so only in the Negative Comment ($p < .01$), 20% Reinforcement ($p < .01$), and No Marble ($p < .01$) conditions, as reflected by significant Verbal X Scale and Objective X Scale interactions. The means for the Verbal X Scale interaction are included in Table 6 and for the Objective X Scale interaction in Table 8. Furthermore, the difference between estimates of past and future performance in the Negative Comment condition was greater in the Negative Comment-20% Reinforcement ($p < .01$) and Negative Comment-No Marble ($p < .01$) groups than in the Negative Comment-80% Reinforcement group, where there was no significant difference between estimates. This fact was indicated by a significant Verbal X Objective X Scale interaction. No other significant effects were found.

Relationship between Estimates of Past and Future Performance and Discrimination Learning Performance

Analyses were performed to examine whether a subject's ratings on the estimate of past performance (for both ages) and estimate of future performance (older subjects only) scales were related to learning performance. The possible ratings on all scales were divided into three subgroups, comprising those who thought they had done or would do poorly, medium, and well, respectively. For the estimate of past and future performance scales for the older subjects, these groups included subjects who had given themselves ratings of 1, 2, or 3 (poor), 4 or 5 (medium), and 6, 7, or 8 (well). For the estimate of performance scale for the younger subjects, the scores were divided into those whose responses the experimenters felt rated a 1 (poor), 2 (medium), or 3 (well). These groupings divided the populations unevenly, but were felt to reflect psychologically meaningful categories. The relationship between scale scores and learning performance was

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 8

Mean Scores of Older Subjects in Each Objective Condition

Omitting No Comment Verbal Condition

on 8-Point Estimate of Past Performance

(and Estimate of Future Performance) Scales

Objective Conditions		
80% Reinforcement	20% Reinforcement	No Marble
5.55 (5.55)	3.40 (4.40)	4.50 (5.20)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

investigated by comparing the mean number of correct responses for the three subgroups on each of the scales.

Estimates of past performance on the experimental games were not related to subsequent performance on the learning task, as indicated by the failure to find any significant differences in mean number of correct responses associated with the estimate of past performance variable at either age. In contrast, estimates of future performance on the learning task of the older subjects were found to be related to subsequent learning performance, in that subjects who expected to do poorly made the most correct responses ($\bar{X}$ = 63.5); those who expected to do well made the fewest correct responses ($\bar{X}$ = 51.7). The performance of those who expected to do medium was between those who expected to do better or worse ($\bar{X}$ = 58.1). Subjects in the low and medium groups were combined to equal in number more nearly the number of subjects in the high group, so that a fair test of differences could be made. The mean number of correct responses of those who scored low and medium combined ($\bar{X}$ = 59.4) was significantly greater than that of those who scored high ($\bar{X}$ = 51.7) (t_{88} = 1.87, $p < .05$), indicating that subjects who expected to do well on the learning task in fact learned more poorly than subjects who made medium or low estimates of future performance. No differences in learning associated with sex were found in relation to the subgroups of the estimate of future performance scale.

Relationships among Estimates of Past and Future Performance, IQ and Children's Social Desirability Scale Scores, and Discrimination Learning Performance of Older Subjects

The older subjects' estimates of past performance on the experimental games were found to be negatively correlated with IQ scores (r = -.26,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$p < .05$), indicating that the more poorly a subject thought he had done on the experimental games, the more intelligent he was likely to be, as measured by IQ scores. Estimates of future performance were also found to be related to IQ, in that subjects who expected to do well on the learning task (high estimate of future performance group) had lower IQ scores ($\bar{X}$ = 109.7) than the medium ($\bar{X}$ = 115.7) and low ($\bar{X}$ = 114.7) groups. When the low and medium groups were combined (to form a group more equal in number to subjects in the high group), it was found that subjects who thought they would do poor or medium on the learning task had significantly higher IQ scores ($\bar{X}$ = 115.4) than subjects who thought that they would do well ($\bar{X}$ = 109.7) ($t_{88} = 1.94$, $p < .05$). No significant differences associated with sex were found among IQ scores in the estimate of future performance groups.

Scores on the Children's Social Desirability scale were examined in relation to estimates of future performance. Higher scores on this measure indicate a greater tendency to give socially desirable responses. The mean social desirability scores for the low, medium, and high estimate of future performance groups were 14.2, 18.4, and 21.2, respectively. These differences were significant ($F_{2,87} = 3.54$, $p < .05$), indicating that subjects in the low estimate of future performance group tended significantly less than subjects in the high estimate of future performance group to give socially desirable responses ($p < .05$). The relationship between estimates of future performance and social desirability scores was also evident in the positive correlation ($r = .25$, $p < .05$) found between the two measures, indicating that children who expected to do well on the learning task also evidenced a greater tendency to give socially desirable responses. A significant difference associated with sex was found for socially desirable

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

response scores in the estimate of future performance groups, as indicated by a significant Sex X Group interaction ($F_{2,54} = 23.7$, $p < .001$) revealed in a 2 X 3 (Sex X Group) analysis of variance. While scores for boys and girls did not differ in the low and middle groups, of subjects who thought they would do well on the learning task (high group), boys tended less to give socially desirable responses ($\bar{X} = 18.8$) than girls ($\bar{X} = 24.1$). No other effects were significant.

Socially desirable response scores were also found to be significantly and negatively related to IQ scores ($r = -.27$, $p < .05$), reflecting the fact that the more intelligent children showed less tendency to give socially desirable responses. To examine further the relationships between socially desirable response scores and IQ scores and learning (number of correct responses), mean scores on the latter two measures as well as for the Children's Social Desirability scale were computed for subjects in the bottom, middle, or top thirds of the socially desirable response scale. The differences in mean scores among the low ($\bar{X} = 10.9$), middle ($\bar{X} = 18.9$), and high ($\bar{X} = 28.4$) socially desirable response groups were significant ($F_{2,87} = 136.06$, $p < .001$). All means were significantly different from each other at $p < .01$. IQ scores for the three socially desirable response groups were also significantly different ($F_{2,87} = 3.93$, $p < .025$), reflecting the fact that the low socially desirable response group had higher IQ scores ($\bar{X} = 118.2$) than the middle ($\bar{X} = 108.9$, $p < .05$) or high ($\bar{X} = 110.3$, $p < .05$) socially desirable response groups. The mean number of correct responses for the three groups was 60.5 (low), 51.8 (middle), and 53.7 (high). Since the means for the middle and high groups were very similar, the number of correct responses in these groups were combined to test the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

difference between learning in the low socially desirable response group and all other subjects. Subjects in the low socially desirable response group made significantly more correct responses ($\bar{X}$ = 60.5) than subjects in the middle and high groups combined ($\bar{X}$ = 52.8) (t_{88} = 1.83, p < .05), indicating that the third of the subjects who tended to give fewer socially desirable responses than the rest of the subjects learned the discrimination task significantly better than the other subjects.

Among the older subjects, IQ scores were found to be significantly related to number of correct responses (r = .22, p < .05), indicating that the more intelligent children learned faster. When socially desirable response scores were partialed out, the relationship between IQ scores and number correct was only slightly greater (r = .25, p < .05), indicating that socially desirable response scores were not a significant influence on the relationship between IQ scores and learning.

No differences associated with sex were found for IQ scores and number of correct responses in the socially desirable response groups. However, overall, boys scored significantly lower ($\bar{X}$ = 17.9) than girls ($\bar{X}$ = 20.8) (t_{88} = 1.86, p < .05) on the Children's Social Desirability scale.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Discussion

The absence of differences in learning associated with age was not unexpected in view of the pretesting which established that the size discrimination task was of equal difficulty using stimuli of different dimensions at the two age levels.

The general expectancy that there would be differences in learning associated with the success-failure manipulations did not receive any direct support, as indicated by the failure to find any overall effects for the verbal and objective manipulations. The absence of any overall effects for the verbal and objective manipulations may be attributable to the difficulty of the task at both ages when the middle-sized stimulus was correct, producing a floor effect which precluded a sensitive test of the effects of the manipulations. Support for this interpretation is obtained from the fact that when subjects for whom the middle-sized stimulus was correct were omitted, an effect for the verbal conditions emerged.

The greater difficulty of the task with the middle-sized stimulus has been reported previously in the literature (Kass & Stevenson, 1961; Steigman & Stevenson, 1960). It may be that a three-choice discrimination task is easier to learn when the large and small stimuli are correct because they are more clearly discriminable from each other than the middle-sized stimulus is from either. Consequently the perceptual demands of the task are greater when the middle-sized stimulus is correct, as reflected in the finding that more errors were made and fewer children learned the discrimination when the middle-sized stimulus was correct.

When the effects of the manipulations on learning were examined omitting those subjects for whom the middle-sized stimulus was correct, a major

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

difference was revealed. Children learned the discrimination task better following preliminary tasks if an adult made no verbal comments about their performance during the preliminary tasks than if the adult verbally evaluated their performance positively or negatively. It should be noted that the present result pertains only to a situation in which success-failure manipulations occur during preliminary tasks and not on the learning task itself. A possible explanation of this finding may be that the evaluative comments by the adult induced in the child a set for receiving feedback about his performance, so that on the learning task, when the adult was silent, the child became anxious in the absence of feedback. His concern over the lack of information from the adult regarding the quality of his performance might then have interfered with his learning. This interpretation is consistent with the formulation that high drive level caused by such factors as anxiety has been found to interfere with the learning of complex tasks (Taylor, 1956).

This finding, that both positive and negative comments by an adult about a child's performance on tasks have a general attenuating effect on an immediately subsequent learning task, appears to differ markedly from the results of previous studies (Butterfield & Zigler, 1965; Kass & Stevenson, 1961; Meid & Zigler, 1971; Steigman & Stevenson, 1960; Stevenson & Zigler, 1958), both in the lack of differences found between the effects of success and failure and in the direction of the effects compared to a control group. The inconsistent findings cannot be explained by differences in the age of subjects used in the various studies, since age was a variable in the present study, and results were comparable for both younger and older children.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The results of the studies by Steigman and Stevenson (1960) and Kass and Stevenson (1961) are discrepant from those of the present study in that the former studies indicated that a success condition facilitated subsequent learning performance more than either failure or control conditions. However, while the manipulation of objective reinforcement in these two studies was comparable to the method in the present study, the verbal manipulations which produced an effect in the present study were considerably stronger than the mere statements of correctness which accompanied the dispensing of marbles in the Steigman and Stevenson and the Kass and Stevenson studies. It may be that the positive and negative comments defining the verbal success and failure conditions in the present study were so strongly worded that they served to attenuate the subsequent learning performance of children in these two conditions by inducing anxiety. This interpretation suggests that the failure subjects may have been extremely anxious beginning the learning task because of the very negative evaluation by the adult of their prior performance, while the success subjects may have been anxious because of the unexpected withdrawal of the intensive positive reinforcement which they had been receiving on the first two games. It may be that there is an optimal strength for both positive and negative comments about a child's performance to best facilitate subsequent learning. One test of this hypothesis would be to vary the strength of positive and negative comments about the child's prior performance and to examine the effect of this variation on learning performance.

Results discrepant with those in the present study were also obtained by Butterfield and Zigler (1965) and Meid and Zigler (1971). These studies employed a procedure very similar to that in the present study and obtained

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

directly opposite results, in that both success and failure conditions were found to facilitate subsequent learning performance more than a control condition. This finding was based on the response measure of trials to criterion; however, trials were defined differently than in the present study.[1] It was possible to recompute the data from the Meid and Zigler study by a method comparable to that in the present study, so that a trial consisted of one presentation of the stimuli and the response measure was number of correct responses through 80 trials. When this was done, no differences in learning associated with condition were found. This result is comparable to the absence of overall effects for the verbal and objective manipulations in the present study. Future investigators of the effects of success and failure on children's learning might well consider which method of measuring learning is most appropriate for their research purposes. They might also design their studies to facilitate comparison to other studies in regard to subject populations, experimental procedure, and response measure employed.

Among children who learned the discrimination in the present study, an interesting sex difference was revealed. Boys learned better with negative than positive or no verbal reinforcement, whereas for girls the situation was reversed. It would seem that girls' learning performance is more subject to disruption by disapproval from an adult than boys'. This finding is consistent with results of other studies indicating that girls'

---

1. A trial was defined by a correct response. However, a correction procedure was used, which consisted of presenting the subject with the same arrangement of stimuli until he made a correct response. Thus a subject who learned the discrimination in a given number of trials might in fact have been presented with the stimuli more times than his score for trials to criterion would indicate.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

academic achievement motivations may be directed more than boys' toward obtaining approval and affection from others (Crandall, 1963). Thus when the girls received disapproval on the preliminary tasks, they may have been less motivated to achieve on a subsequent task, while the boys experienced an increase in motivation. This pattern of sex differences is also consistent with the finding of Crandall and Rabson (1960) that failure on a task motivated boys ages six to eight to try again to succeed on it when given another chance, rather than to repeat a task that they had previously performed successfully. Girls, on the other hand, more often chose to repeat earlier successes than failures. In addition, boys in a situation with female experimenters may generally be more motivated to perform well than girls by a desire to please an adult of the opposite sex. This interpretation is consistent with the literature demonstrating that women are more effective reinforcers with boys of this age than with girls, whereas men tend to be more effective with girls than with boys (Stevenson, 1961, 1965). However, most studies which have demonstrated this cross-sex effect have employed relatively simple motor and persistence tasks. Thus it would be interesting to investigate in further studies of learning and problem solving whether the sex of the experimenter relates to the effects of success and failure experiences on boys' and girls' subsequent learning by including experimenters of both sexes.

The most striking developmental finding of the present study concerned children's feelings about their past performance on the experimental games. In evaluating their past performance, the younger children were sensitive only to the verbal manipulation and did not respond to the objective manipulation, whereas the older children responded to both. Six-year-old

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

children were more likely to define success and failure by what an adult told them rather than by forming their own evaluation based on objective evidence. This result contributes further evidence that adult praise and approval are very important to young children and are instrumental in defining the positive and negative aspects of their experiences. It is consistent with the findings that young children are more responsive to adult praise than to knowledge of correctness on a task than older children (Harter, 1967; Stevenson, 1961, 1965); that they perceive the outcome of events as being externally rather than internally controlled more than older children (Cromwell, 1963); and that young children demonstrate a greater need for adult approval by making more socially desirable responses than older children (Crandall et al., 1965). It is interesting to note that for the younger children, the negative comments had a greater effect on feelings about past performance than the positive comments, relative to the control group, indicating that the younger children were more sensitive to disapproval from the adult than approval.

The finding that the younger girls rated themselves lower overall than the boys on how well they thought they had done on the games may demonstrate a general trend for girls of this age to feel inferior to boys. This finding is consistent with the socialization literature on sex-role development (Maccoby, 1966). For example, investigators have found a similar tendency for girls to estimate their future performance worse than boys. Crandall, Katovsky, and Preston (1962), using children of above-average intelligence in first, second, and third grades, found that the more intelligent girls expected to fail on a new task, but the more intelligent boys expected to succeed, in comparison to children of less

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

intelligence. At the older age of 10 years, the differences between boys and girls in ratings on estimate of past performance disappeared in the present study. It may be that by this age, girls generally feel better about themselves in comparison to boys than they did when younger, having learned that they can achieve as well as boys, particularly in academic and intellectual activities. For this reason they might no longer rate themselves inferior to boys.

The findings for the older children that both experimental manipulations influenced their feelings about their past performance indicated that by the age of 10, children still seem to be sensitive to verbal comments from an adult about their performance, but they have also gained the ability to form evaluations of their performance from relevant objective evidence. Furthermore, they use the two methods of evaluation in combination, and when one source of information conflicts with the other, they allow the more favorable evaluation, regardless of source, to predominate, although the negative evaluation also influences their estimates. Thus when receiving both positive and negative verbal and objective evidence about their performance at the same time, their estimates fell between those formed when both types of evidence were either positive or negative. However, evaluations based on conflicting evidence tended to be higher than the average between those formed from doubly positive or doubly negative sources.

The major effect of the verbal manipulation on the older children was like that for the younger children, i.e., negative comments served to lower estimates of past performance relative to success and control conditions, while the latter two conditions did not differ. This effect was more

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

pronounced under conditions of objective failure than of objective success, such that when the children were winning few marbles, the adult's evaluations of their performance had greater differential effects than when the children were winning lots of marbles. This finding may be related to the hypothesis of Crandall et al. (1965) that younger children tend to try to appear more socially desirable than older children do because they are more dependent on adult approval due to their instrumental inadequacy and emotional dependency. Under conditions in which the older children were apparently instrumentally incompetent, as evidenced by their failure to win many marbles, they demonstrated a greater differential responsiveness to adult verbal reinforcement. Thus when feeling relatively incompetent, as though they were younger, they tended to respond to the adult's verbal reinforcement more than they would otherwise, as younger children might. In the absence of any objective indicators of success and failure, i.e., when the children were not playing the games for marbles, they showed their sensitivity to the positive and negative verbal reinforcement by estimating their past performance as poorer when the adult made negative comments.

The older children proved to be sensitive to the objective indicators of success and failure in a more complex way than to the verbal reinforcement, in that both success and failure defined objectively influenced their feelings about their past performance, relative to the control group. Winning many marbles raised estimates of past performance and winning few marbles lowered them, relative to the estimates in the control condition. The children were more sensitive in this way to differences in the number of marbles won when the adult made negative comments about their performance than when the adult was praising them, for when the adult praised their

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

performance, there was no difference in their estimates whether they had won many marbles or had not played the games for marbles. It would seem that under conditions of adult disapproval, in contrast to the situation of objectively defined failure, the older children, being more autonomous and less dependent on adult evaluations than younger children, became more motivated to discount the adult's unfavorable judgment of their performance by relying on objective evidence. Under conditions of adult approval, they were not so motivated to attend to objective bases of evaluation. In the absence of any comments from the adult about their performance, the older children demonstrated their sensitivity solely to the objective indicators of success and failure by estimating their performance as better when they had won many marbles than when they had won few.

The effects of the verbal and objective manipulations on the older children's estimates of future performance on the learning task were similar to those on their estimates of past performance on the experimental games. However, estimates of past performance were generally lower overall than estimates of future performance. Furthermore, the manipulations differentially affected estimates of past performance more than estimates of future performance. These findings are not surprising, given that the manipulations occurred on and were directed toward the children's "past" performance.

The groups in which the differences between the effects of the preliminary conditions on the estimates of past and future performance were greatest were those groups having the greatest failure implications, in which the adult both made negative comments about the children's performance and the children won few marbles or did not play for marbles. The greater

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

differential effect of negative than positive reinforcement, whether defined verbally or objectively, in comparison to a control group, may be attributable to the generally high expectancies of success of white, middle-class children in a variety of achievement situations. These children, not having the histories of repeated failure which have been found to promote generally low expectations of success (Gruen & Zigler, 1968; Stevenson & Zigler, 1958; Turnure & Zigler, 1964; Zigler, 1966) would respond more to indications of failure than of success, because failure would be discrepant with their general expectancy of doing well. Success would not then have so great an effect, since it would be consistent with their expectations. This effect was true even for the 6-year-olds on their estimates of past performance.

An apparent paradox emerged in the relationship between the older children's estimates of future performance on the learning task and how well they actually learned: Children who thought that they would do well on the learning task in fact learned more poorly than children who had low or moderate estimates of future performance. However, the relationships among estimates of future performance, learning, the tendency to give socially desirable responses, and intelligence help to clarify this pattern. Nevertheless, while these correlational data describe an interesting constellation of traits which occurred consistently in the older children, they do not provide a causal explanation of the pattern which emerged.

In view of the positive relationship found between estimates of future performance and scores on the Children's Social Desirability scale, it seems possible that the estimate of future performance measure may have reflected a more global tendency to attempt to appear socially desirable.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Thus children who said that they thought that they would do well on the learning task may have said so motivated by a greater need to obtain adult approval by appearing socially desirable than those children who were less certain of their good future performance. The attempt to appear socially desirable was manifested in saying what they thought would please the adult, i.e., that they would be competent on the learning task, or at least attempt to do well. Furthermore, concern over the adult's approval or disapproval may have interfered with these children's learning, so that consequently they learned more poorly than the other children. This interpretation is substantiated by the finding that the third of the older children scoring lowest on tendency to give socially desirable responses learned significantly better than the other children. This finding is consistent with that of Crowne, Holland, and Conn (1968) that fifth- and sixth-grade children high in need for adult approval performed significantly worse on a discrimination learning task than children low in need for approval, as measured by the Children's Social Desirability scale. The hypothesis suggested to explain this result was that anxious arousal of the approval-dependent children when faced with an achievement situation in which they might fail and lose approval interfered with selective attentional processes necessary to learn the discrimination.

The negative relationship between the tendency to give socially desirable responses and IQ is consistent with the finding of Crandall et al. (1965) that children who score high on the Children's Social Desirability scale tend to have lower IQs. It may be that the better learning performance of children who scored low on the Children's Social Desirability scale was related to their greater intelligence. This interpretation gains

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

some support from the positive relationship found between IQ and learning. Furthermore, the children who had high estimates of future performance and learned more poorly than the other children also had lower IQs than the other children. However, the relationship found between IQ and learning was weak. Also, the differences in IQ between the children low in tendency to give socially desirable responses and the other children and children high on estimate of future performance and the other children were not large. Thus, evidence supporting the relationship between IQ and differences in learning associated with estimates of future performance and the Children's Social Desirability scale, while consistent, is not compelling.

It is apparent from the complicated relationships among learning, estimates of future performance, tendency to give socially desirable responses, and intelligence in the present study that additional research is necessary to ferret out the intricacies among them. Furthermore, one of the more important implications of the present study is that future research in the area of the effects of success and failure on children's learning should concern intervening process variables, such as how children feel in regard to success-failure manipulations employed, as well as the additional measures such as IQ and tendency to give socially desirable responses, which in turn help to clarify how children will evaluate their performance. Attention to such intervening measures might lessen the confusion resulting from apparently different effects of success and failure on children's learning in different studies. Quite possibly the variable to examine to predict learning performance best is not the prior experimental success or failure or neutral experience per se, however operationally defined, but the child's feelings about how he is going to do on the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

subsequent task, his estimate of future performance. Consequently, more sensitive measures of children's evaluations of their performance are necessary. Since children are the focus of child development research, it should be their feelings which define effective success-failure manipulations, which then might better predict their learning performance.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

References

Butterfield, E. C., & Zigler, E. The effects of success and failure on the discrimination learning of normal and retarded children. Journal of Abnormal Psychology, 1965, 70, 25-31.

Crandall, V. C., Crandall, V. J., & Katovsky, W. A children's social desirability questionnaire. Journal of Consulting Psychology, 1965, 29, 27-36.

Crandall, V. J. Achievement. In H. W. Stevenson (Ed.), Child psychology. Part I. Chicago: National Society for the Study of Education, 1963.

Crandall, V. J., Katovsky, W., & Preston, A. Motivational and ability determinants of young children's intellectual achievement behaviors. Child Development, 1962, 33, 643-661.

Crandall, V. J., & Rabson, A. Children's repetition choices in an intellectual achievement situation following success and failure. Journal of Genetic Psychology, 1960, 97, 161-168.

Cromwell, R. L. A social learning approach to mental retardation. In N. R. Ellis (Ed.), Handbook of mental retardation. New York: McGraw-Hill, 1963.

Crowne, D. P., Holland, C. H., & Conn, L. K. Personality factors in discrimination learning in children. Journal of Personality and Social Psychology, 1968, 10, 420-430.

Gruen, G., & Zigler, E. Expectancy of success and the probability learning of middle-class, lower-class, and retarded children. Journal of Abnormal Psychology, 1968, 73, 343-352.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Harter, S. Mental age, IQ, and motivational factors in the discrimination learning set performance of normal and retarded children. Journal of Experimental Child Psychology, 1967, 5, 123-141.

Kass, N., & Stevenson, H. W. The effect of pretraining reinforcement conditions on learning by normal and retarded children. American Journal of Mental Deficiency, 1961, 66, 76-80.

Maccoby, E. E. (Ed.) The development of sex differences. Stanford, Calif.: Stanford University Press, 1966.

Meid, E., & Zigler, E. The effects of actual and implied success and failure on the discrimination learning of normal children. Unpublished manuscript, Yale University, 1971.

Steigman, M. J., & Stevenson, H. W. The effects of pretraining reinforcement schedules on children's learning. Child Development, 1960, 31, 53-58.

Stevenson, H. W. Social reinforcement with children as a function of CA, sex of E, and sex of S. Journal of Abnormal Psychology, 1961, 63, 147-154.

Stevenson, H. W. Social reinforcement of children's behavior. In L. P. Lipsitt & C. C. Spiker (Eds.), Advances in child development and behavior. Vol. 2. New York: Academic Press, 1965.

Stevenson, H. W., & Zigler, E. Probability learning in children. Journal of Experimental Psychology, 1958, 56, 185-192.

Taylor, J. A. Drive theory and manifest anxiety. Psychological Bulletin, 1956, 53, 303-318.

Turnure, J., & Zigler, E. Outer-directedness in the problem solving of normal and retarded children. Journal of Abnormal and Social Psychology, 1964, 69, 427-436.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Zigler, E. Research on personality structure in the retardate. In N. R. Ellis (Ed.), International review of research in mental retardation. Vol. 1. New York: Academic Press, 1966.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Appendix A

Questions and Bar Scales

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Estimate of Past Performance

Following the second preliminary games (or after 7 minutes for subjects in the No Comment-No Marble group, the experimenter asked, "Were you doing anything special with the cards and pictures or not?" Depending on the subject's response, the experimenter asked further, "What were you doing?" After the subject had had an opportunity to describe his strategy on the games (or what he had been doing with the materials), he was asked, "How well do you think you did? Why?" Subjects in the No Comment-No Marble group were not asked this question unless their response to the first questions indicated that they had elaborated the preliminary activity into a game or something similar. (In the majority of cases, they had not.)

Then the estimate of past performance bar scale was introduced to all subjects except those in the No Comment-No Marble group. It consisted of eight vertical bars, approximately 1/4-inch (6.4mm) wide, increasing in length from left to right. Beneath each bar was a small stick figure which could represent a boy or a girl. Beneath each stick figure was a small square box which the subject could mark with a pencil. Instructions for this bar scale were:

> These bars show how well some other boys (girls) did on the games. The longer the bar, the better the boy (girl) played. Show me the one who did the best. . . . Right, because he (she) has the longest bar. Now show me the one who did the worst. . . . Right, because he (she) has the shortest bar. Now show me a boy (girl) who did medium, not too good but not too bad. . . . Right, because he (she) has a medium-sized bar. Which one shows

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

how well you think you did? Okay, put a check in the box underneath. Why did you pick that one?

After marking the bar scale, subjects in the 80% Reinforcement and 20% Reinforcement conditions were asked, "Do you think you won enough marbles to get the prize? Why?" Subjects in the Positive Comment-No Marble and Negative Comment-No Marble groups were asked, "Do you think you played the games well enough to get the prize? Why?"

Apart from the first two questions regarding what they had been doing with the game materials during the 7 minutes of preliminary activity, subjects in the No Comment-No Marble group were asked nothing concerning estimate of past performance.

Estimate of Future Performance

After the subject had responded to the last question about his estimate of past performance, the experimenter introduced the estimate of future performance bar scale. It was identical to the one for estimate of past performance. This time the experimenter said:

> These bars show how well some other boys (girls) did on the third game. Again, the longer the bar, the better the boy (girl) played, and the shorter the bar, the worse he (she) played. Which bar shows how well you think you're going to do on the third game? Okay, put a check in the box underneath. Why did you pick that one?

Subjects in the No Comment-No Marble group also were presented with the estimate of future performance scale. To them, the experimenter said:

> These bars show how well some other boys (girls) did on the game. The longer the bar, the better the boy (girl) played.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Show me the one who did the best. . . . Right, because he (she) has the longest bar. Now show me the one who did the worst. . . . Right, because he (she) has the shortest bar. Now show me a boy (girl) who did medium, not too good but not too bad. . . . Right, because he (she) has a medium-sized bar. Which one shows how well you think you're going to do? Okay, put a check in the box underneath. Why did you pick that one?

Learning Task

After the subject had reached the criterion of five consecutive correct responses or had been presented with the stimuli for 80 trials, he was asked the following questions: "What did you think about Find-the-Marble? What did you think you had to do in it? How did you decide which one to pick each time?" These questions were asked of all subjects to see if those who reached criterion could verbalize the solution to the problem, and if those who did not reach criterion were approaching a solution. (If a subject verbalized the solution to any one of the questions, no further questions were asked.)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Appendix B

Results Based on Trials to Criterion

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table A

Analysis of Variance of Trials to Criterion for Learners Only by Objective Conditions, Verbal Conditions, Age, and Sex

Source	df	MS	F
Objective Conditions (A)	2	282.48	<1
Verbal Conditions (B)	2	242.17	<1
Age (C)	1	1935.66	4.44*
Sex (D)	1	604.80	1.39
A X B	4	106.81	<1
A X C	2	464.42	1.07
A X D	2	408.80	<1
B X C	2	202.66	<1
B X D	2	1723.90	3.96**
C X D	1	53.95	<1
A X B X C	4	206.90	<1
A X B X D	4	300.93	<1
A X C X D	2	63.08	<1
B X C X D	2	244.82	<1
A X B X C X D	4	363.10	<1
Error (W)	96	435.64	

*$p < .05$
**$p < .025$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table B

Analysis of Variance of Trials to Criterion for Learners Only by Objective Conditions, Verbal Conditions, Age, and Experimenter

Source	df	MS	F
Objective Conditions (A)	2	741.13	1.80
Verbal Conditions (B)	2	314.22	<1
Age (C)	1	1020.37	2.48
Experimenter (D)	1	544.60	1.32
A X B	4	120.36	<1
A X C	2	384.47	<1
A X D	2	911.67	2.21
B X C	2	423.77	1.03
B X D	2	1.78	<1
C X D	1	2834.74	6.88*
A X B X C	4	172.00	<1
A X B X D	4	396.37	<1
A X C X D	2	397.54	<1
B X C X D	2	388.36	<1
A X B X C X D	4	369.76	<1
Error (W)	96	411.87	

*$p < .025$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table C

Mean Trials to Criterion for Learners Only

at Each Age by Sex in the Verbal Conditions

Age	Positive Comment		Negative Comment		No Comment	
	Male	Female	Male	Female	Male	Female
6	40.5	40.4	42.4	44.3	34.2	28.9
10	34.1	22.9	18.1	34.2	28.0	30.3

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table D

Mean Trials to Criterion for Learners Only at Each Age for Each Experimenter

Age	Experimenter 1	Experimenter 2
6	40.32	27.80
10	24.88	30.54

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Appendix C

Summary F Tables of Analyses of Variance of Estimate of Past and Future Performance Scales

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table A

Analysis of Variance of Scores on Estimate of Past Performance Scale for Younger Subjects by Verbal Conditions, Objective (Omitting No Marble) Conditions, and Sex

Source	df	MS	F
Verbal Conditions (A)	2	9.02	29.10***
Objective Conditions (B)	1	.17	<1
Sex (C)	1	1.51	4.87**
A X B	2	.91	2.94*
A X C	2	.72	2.32
B X C	1	.17	<1
A X B X C	2	.34	1.10
Error (W)	45	.31	

*$p < .075$
**$p < .05$
***$p < .001$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table B

Analysis of Variance of Scores on Estimate of Past Performance Scale for Younger Subjects by Verbal (Omitting No Comment) Conditions, Objective Conditions, and Sex

Source	df	MS	F
Verbal Conditions (A)	1	22.46	66.06*
Objective Conditions (B)	2	.34	<1
Sex (C)	1	1.06	3.12
A X B	2	.02	<1
A X C	1	.68	2.00
B X C	2	.59	1.74
A X B X C	2	.40	1.18
Error (W)	46	.34	

*$p < .001$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table C

Analysis of Variance of Scores on Estimate of Past Performance Scale for Older Subjects by Verbal Conditions and Objective (Omitting No Marble) Conditions

Source	df	MS	F
Verbal Conditions (A)	2	30.22	26.05**
Objective Conditions (B)	1	58.02	50.02**
A X B	2	6.01	5.19*
Error (W)	54	1.16	

*$p < .01$
**$p < .001$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table D

Analysis of Variance of Scores on Estimate of Past Performance Scale for Older Subjects by Verbal Conditions, Objective (Omitting No Marble) Conditions, and Sex

Source	df	MS	F
Verbal Conditions (A)	2	30.22	27.47**
Objective Conditions (B)	1	58.02	52.74**
Sex (C)	1	1.34	1.22
A X B	2	6.02	5.47*
A X C	2	1.34	1.22
B X C	1	2.02	1.84
A X B X C	2	1.82	1.65
Error (W)	48	1.10	

*$p < .01$
**$p < .001$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table E

Analysis of Variance of Scores on Estimate of Past Performance Scale for Older Subjects by Verbal (Omitting No Comment) Conditions and Objective Conditions

Source	df	MS	F
Verbal Conditions (A)	1	98.82	99.82**
Objective Conditions (B)	2	23.12	23.35**
A X B	2	6.22	6.28*
Error (W)	54	.99	

*$p < .005$
**$p < .001$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table F

Analysis of Variance of Scores on Estimate of Past Performance Scale for Older Subjects by Verbal (Omitting No Comment) Conditions, Objective Conditions, and Sex

Source	df	MS	F
Verbal Conditions (A)	1	98.82	107.41**
Objective Conditions (B)	2	23.12	25.13**
Sex (C)	1	1.34	1.46
A X B	2	6.22	6.76*
A X C	1	1.34	1.46
B X C	2	2.44	2.66
A X B X C	2	.94	1.02
Error (W)	48	.92	

*$p < .005$
**$p < .001$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table G

Analysis of Variance of Scores on Estimate of Future Performance Scale for Older Subjects by Verbal Conditions and Objective Conditions

Source	df	MS	F
Verbal Conditions (A)	2	18.30	12.12***
Objective Conditions (B)	2	10.30	6.82**
A X B	4	6.34	4.21*
Error (W)	81	1.51	

*$p < .025$
**$p < .005$
***$p < .001$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table H

Analysis of Variance of Scores on Estimate of Future Performance Scale for Older Subjects by Verbal Conditions, Objective Conditions, and Sex

Source	df	MS	F
Verbal Conditions (A)	2	18.30	11.51**
Objective Conditions (B)	2	10.30	6.48*
Sex (C)	1	1.88	1.48
A X B	4	6.34	3.99**
A X C	2	1.68	1.06
B X C	2	.28	< 1
A X B X C	4	.43	< 1
Error (W)	72	1.59	

*$p < .005$
**$p < .001$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table I

Analysis of Variance of Scores on Estimate of

Past and Future Performance Scales by

Verbal Conditions and Objective (Omitting No Marble) Conditions

Source	df	MS	F
Between Subjects[a]			
Verbal Conditions (A)	2	41.31	19.77**
Objective Conditions (B)	1	70.53	33.75**
A X B	2	17.31	8.28**
Error (B)	54	20.87	
Within Subjects			
Scale (T)	1	7.50	16.67**
A X T	2	2.03	4.51*
B X T	1	5.63	12.51**
A X B X T	2	.76	1.69
Error (W)	54	.45	

[a]The analysis for Between Subjects, based on the mean of the scores for both scales together, does not pertain to the purpose of the analysis.

*$p < .025$
**$p < .001$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table J

Analysis of Variance of Scores on Estimate of Past and Future Performance Scales for Older Subjects by Verbal Conditions, Objective (Omitting No Marble) Conditions, and Sex

Source	df	MS	F
Between Subjects[a]			
Verbal Conditions (A)	2	41.31	19.39***
Objective Conditions (B)	1	70.53	33.11***
Sex (C)	1	2.13	1.00
A X B	2	17.31	8.13
A X C	2	1.61	<1
B X C	1	2.13	1.00
A X B X C	2	1.41	<1
Error (B)	48	2.13	
Within Subjects			
Scale (T)	1	7.50	17.04***
A X T	2	2.03	4.60*
B X T	1	5.63	12.73***
C X T	1	.03	<1
A X B X T	2	.76	1.73
A X C X T	2	.86	1.95
B X C X T	1	.30	<1
A X B X C X T	2	.53	1.20
Error (W)	48	.44	

[a]See footnote, Table I.

*$p < .025$
**$p < .005$
***$p < .001$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table K

Analysis of Variance of Scores on Estimate of Past and Future Performance Scales for Older Subjects by Verbal (Omitting No Comment) Conditions and Objective Conditions

Source	df	MS	F
Between Subjects[a]			
Verbal Conditions (A)	1	112.13	55.72***
Objective Conditions (B)	2	27.43	13.65***
A X B	2	16.93	8.42***
Error (B)	54	2.01	
Within Subjects			
Scale (T)	1	9.63	20.06***
A X T	1	12.03	25.06***
B X T	2	2.63	5.48**
A X B X T	2	1.63	3.40*
Error (W)	54	.48	

[a]See footnote, Table I.

*p < .05
**p < .01
***p < .001

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table L

Analysis of Variance of Scores on Estimate of Past and Future Performance Scales for Older Subjects by Verbal (Omitting No Comment) Conditions, Objective Conditions, and Sex

Source	df	MS	F
Between Subjects[a]			
Verbal Conditions (A)	1	112.13	54.17****
Objective Conditions (B)	2	27.43	13.25****
Sex (C)	1	.83	<1
A X B	2	16.93	8.18***
A X C	1	1.63	<1
B X C	2	2.23	1.08
A X B X C	2	1.23	<1
Error (B)	48	2.07	
Within Subjects			
Scale (T)	1	9.63	19.65****
A X T	1	12.03	24.55****
B X T	2	2.63	5.37**
C X T	1	.53	1.08
A X B X T	2	1.63	3.33*
A X C X T	1	.13	<1
B X C X T	2	.53	1.08
A X B X C X T	2	.23	<1
Error (W)	48	.49	

[a]See footnote, Table I.

*p < .05
**p < .01
***p < .005
****p < .001

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.