

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Vicarious Reinforcement of Cognitive Skill

A dissertation presented in partial
fulfilment of the requirements for the
degree of Doctor of Philosophy
in Education at Massey University

by

Mildred Sharon **Sass**

1976 (Awarded)

ABSTRACT

This study was designed to explore direct and vicarious reinforcement of cognitive skill. Cognitive skill, in this case, was defined by specific types of answering behaviour.

In developmental stages of the study, a taxonomy of educational objectives was found unsatisfactory for classifying answers given by subjects. For this reason, a classification scheme was developed for use in the present study. This scheme supplied six categories of cognitive skill. Subjects' answers were classified according to which of the six cognitive skills they represented.

Developmental work also explored a variety of types of questions (completion, multiple-choice, and open-ended) in an attempt to determine which were efficient at eliciting answers representing a variety of cognitive skills. Open-ended questions appeared most satisfactory because they allowed free manifestation of a variety of types of answers rather than predetermining the cognitive skill represented by answers, as did completion and multiple-choice items.

However, with open-ended questions, there was an accompanying difficulty of eliciting answers which represented specific categories of response so that consistent stimuli could be provided for observing subjects. Since some categories had low operant levels, a videotape was used to provide standardized and consistent stimuli for observers. The modelling situation did not permit the exploration of direct reinforcement so the emphasis of the study turned to vicarious reinforcement.

In an experimental situation, subjects (girls aged 15 to 16 years) watched a model giving "answers" which corresponded to all six categories of the classification scheme. Some of the model's "answers" were praised and others were not. Subsequent measures of the observing subjects' responses tested the effect of vicarious praise. Measures were also taken to determine the effect of

awareness of contingencies of reinforcement on subjects' responses. Furthermore, alterations in the experimental design were used to explore the effects of increased intensity and relative position of praise in the modelled sequence.

Exact probabilities were used to analyze the data where possible. However, where there was more than one degree of freedom, a chi square statistic was used to determine the significance of results.

The results showed that vicarious praise did influence the types of answers given by observing subjects. That is, the praise was an effective vicarious reinforcer. However, the order of the model's "answers" and awareness on the part of subjects were unrelated to responses given by subjects. Moreover, increased intensity and change in position of the praise did not produce greater treatment effects. Alternate methods of classifying the data were explored but not found to be more efficient than the original scheme at detecting a treatment effect. Also, an unsuccessful attempt was made to identify characteristics of subjects which affected their answering behaviour.

The results are discussed as they relate to both education and psychology, and implication for both fields are considered.

ACKNOWLEDGEMENTS

In completing this document, I am indebted to the numerous volunteers who cooperated at all stages of the investigation. Without them, the project could not have been completed.

I am also grateful to the following individuals who have given invaluable assistance in various stages of this thesis:

Prof. R.S. Adams, for his initial encouragement and continuing guidance;

Prof. C.G.N. Hill and Dr. D. McAlpine, for sound suggestions and support:

Dr. B. Ryan and Mr. E. Archer, for theoretical advice and discussions;

Mr. M. Warn, for hours of videotape;

Mr. G. Clayton and Mr. E. Graham, for remarkable patience in the arduous task of coding responses;

Mr. D. Hoffman and Encyclopaedia Britannica, for loan of the film Reflections on Time;

Dr. R.E. Sass, for untold patience, understanding, and support for so many years.

The success of this thesis is due to efforts of those mentioned above. However, I must take full credit for any shortcomings.

TABLE OF CONTENTS

| | Page |
|--|------|
| ACKNOWLEDGEMENTS | iv |
| TABLE OF CONTENTS | v |
| LIST OF TABLES | vi |
| LIST OF FIGURES | ix |
| INTRODUCTION | 1 |
| CHAPTER | |
| 1. Background and Related Research | 4 |
| 2. Procedure - Design Development | 25 |
| 3. Procedure - Experimental Design | 55 |
| 4. Results | 87 |
| 5. Discussion and Conclusions | 129 |
| APPENDICES | |
| A Short-Answer, Multiple-Choice and Oral Items Used in Developmental Work | 149 |
| B Transcript of Open-ended Questioning of One Subject | 175 |
| C Alternate Definitions of the Six Categories of the Taxonomy of Educational Objectives | 185 |
| D Original and Final Form of the Coding Schedule and Example Response Coding Used to Train Judges | 189 |
| E "Answers" as Classified Using the Taxonomy and the Coding Schedule | 196 |
| F Scripts for Videotapes and Experimental Sessions | 203 |
| G Transcript of Questioning of a Subject after Experimental Treatment | 215 |
| H Complete List of Codings for Personal and Empathetic Questions | 217 |
| I Computer Programmes and Calculations | 220 |
| BIBLIOGRAPHY | 227 |

LIST OF TABLES

| | | Page |
|----|---|------|
| 1 | Summary of Research Dealing with Direct Reinforcement and Vicarious Learning | 22 |
| 2 | Behavioural Terms for Each Level of the Taxonomy | 31 |
| 3 | Examples of Short-Answer Completion Items arranged Hierachically According to the Taxonomy of Educational Objectives | 32 |
| 4 | Number of Correct Completion Answers at Each of the Six Levels of the Taxonomy with a Maximum Score of 4 at Each Level | 35 |
| 5 | Examples of Multiple-Choice Items Arranged According to their Cognitive Level of the Taxonomy of Educational Objectives | 37 |
| 6 | Number of Correct Multiple-Choice Answers at Each of the Six Levels of the Taxonomy with a Maximum Score of 3 at Each Level | 38 |
| 7 | Number of Correct Answers at Each of the Six Levels of the Taxonomy with a Maximum Score of 1 at Each Level - Short Oral Test | 40 |
| 8 | Number of Correct Answers at Each of the Six Levels of the Taxonomy with a Maximum Score of 2 at Each Level - Long Oral Test | 40 |
| 9 | Examples of Open-ended and Short-Answer Questions Asked of One Subject | 43 |
| 10 | Answers to Open-ended and Short-Answer Questions from One Subject, Classified According to the Taxonomy by One Judge | 45 |
| 11 | Percent of Total Agreement of Classifications Between Four Independent Coders | 47 |
| 12 | Order of Response Categories on Baseline Videotapes | 62 |
| 13 | Order of Response Categories on Treatment Videotapes | 64 |
| 14 | Long Phrases of Praise Used for Each Category of the Coding Schedule | 66 |

LIST OF TABLES cont.

| | Page |
|---|------|
| 15 Baseline and Treatment Videotapes Administered to Six Volunteers | 70 |
| 16 Baseline and Treatment Videotapes Randomly Assigned to Subjects | 79 |
| 17 Percentage of Disagreements in Coding Involving Six Categories over Two Questions | 83 |
| 18 Distribution of Baseline Response Categories | 88 |
| 19 Distribution of Position in Baseline Videotape of Baseline Response Categories | 91 |
| 20 Distribution of Position in Treatment Videotape of Initial Response Categories to the Personal Question | 92 |
| 21 Distribution of Position in Treatment Videotape of Representative Response Categories to the Personal Question | 93 |
| 22 Frequencies of Subjects Matching and Not Matching the Praised Model Response with the Initial Response Category | 96 |
| 23 Chance Probability of at Least X Matches with Praised Model Response When $N=30$ and $p=.1667$ | 97 |
| 24 Frequencies of Subjects Matching and Not Matching the Praised Model Response using Representative Response Category | 98 |
| 25 Frequency of Subjects Matching and Not Matching the Praised Model Response using Response Category to the Empathetic Question as Criterion | 99 |
| 26 Chance Probability of at Least X Matches with Praised Model Response When $N=13$, $p=.1667$ | 100 |
| 27 Frequency of Subjects Matching and Not Matching the Category of Response Following the Praise, using Initial Response Category to the Personal Question as Criterion | 102 |
| 28 Frequency of Subjects Matching and Not Matching the Praised Model Response When Employing More Intense Praise | 105 |

LIST OF TABLES cont.

| | Page | |
|----|---|-----|
| 29 | Frequency of Subjects Matching and Not Matching the Praised Model Response When Employing Praise at the End of the Modelled Sequence | 107 |
| 30 | Awareness X Matching Praised Model Response (Representative Response Category) for the Personal Question | 109 |
| 31 | Intensity of Praise of Aware Subjects X Matching Praised Model Response (Representative Category) for the Personal Question | 110 |
| 32 | Intensity of Praise X Awareness for the Personal Question | 111 |
| 33 | Null Hypotheses, Variables and Decisions used to Investigate the Major Hypotheses | 114 |
| 34 | Number of Subjects Matching and Not Matching the Praised Model Response using a (1,2,3) and (4,5) Coding | 118 |
| 35 | Binomial Probabilities for $N=15$, $P=.5216$ | 119 |
| 36 | Number of Subjects Matching and Not Matching the Praised Model Response using a (1) and (2,3) and (4,5) Coding | 120 |
| 37 | Binomial Probabilities for $N=19$, $p=.352$ | 121 |
| 38 | Code 0, 1, and 2 X Answering the Empathetic Question | 123 |
| 39 | Mover vs. Non-Mover X Answering the Empathetic Question | 124 |
| 40 | Code 1 or 2 X Answering the Empathetic Question | 125 |
| 41 | Frequency of Subjects Responding with Same and Different Response Category for Personal Question (Representative Response Category) and the Empathetic Question | 126 |
| 42 | Null Hypotheses, Variables and Decisions Used to Investigate the Minor Hypotheses | 128 |

LIST OF FIGURES

| FIGURE | | Page |
|--------|---|------|
| 1 | Conditions under which a Treatment Effect may be Detected | 74 |
| 2 | Power Curves for $N=18, 24, 30$ | 76 |

Introduction

The Problem

This study is concerned with the effects of reinforcers on cognitive skills. More specifically, it concentrates on "answering behaviours" commonly found in classrooms. To this end it focuses on the "teacher question-student answer" dialogue - a dialogue usually involving one student who actively participates and may be praised for a particular type of answer, and other students who observe this dialogue and who thus may observe the praise given to the participant.

In classrooms, teachers attempt to strengthen a multitude of behaviour by praising or otherwise rewarding students. When teachers offer praise or other stimuli that have the potential of being reinforcers, they are working on the assumption that the effect will be to strengthen the desired behaviour so that the behaviour will re-occur subsequently. It seems that teachers also operate on the assumption that the other students who observe the behaviour will learn vicariously and thus display the desired behaviour.

Research has supported these assumptions to some extent. For example, reinforcers applied directly to the participants have been shown to influence common classroom behaviour such as being on time (Hall, Cristler, Cranston, and Tucker, 1970) and attending to work (Zimmerman, Zimmerman, and Russell, 1969). Moreover, vicarious reinforcers have been shown to influence information-seeking (Henderson and Garcia, 1973) and aggressive acts (Bandura, 1965). Indeed, current research on vicarious and direct reinforcers show that a great variety of behaviour related to the operation of classrooms can be strengthened or increased by the use of contingent reinforcers. Nevertheless, most of that research has been preoccupied with discipline oriented behaviours. Approaches demonstrating that cognitive skills and academic behaviours, such as asking questions and correctly completing mathematics problems, are affected by reinforcers have been relatively infrequent. More specifically related to the present investigation, whether teacher praise strengthens specified answering behaviour in

participating and observing students is currently unknown; if it has been investigated it has not been reported.

Historically, reinforcement theory evolved from infrahuman studies in which behaviour was found to be governed by reinforcing stimuli. Human studies developed subsequently, prompted by a desire to control aberrant behaviour such as phobias, deviancy, alcoholism, and homosexuality. Reinforcement came to be seen as a therapeutic technique for aiding abnormal individuals and as a teaching technique for retardates (See for example Bandura, 1969; Neisworth and Smith, 1973). The studies dealing with infrahuman and abnormal human behaviour together demonstrate the wide range of behaviour which could be successfully altered or controlled through the reinforcement of desired acts. Even so, as far as education is concerned, the studies give a narrow view of the types of behaviour which might be governed by contingent reinforcers.

The domain of behaviour of interest to experimental psychology has been quite different from the domain of behaviour of interest to clinical psychology, both of which differ from the domain of behaviour of interest to educational psychology. The present study thus addresses itself to two audiences - to educators who are concerned with the attainment of cognitive goals and to psychologists who wish to test the domain of behaviour governed by reinforcement theory.

Thus this study was an attempt to determine if cognitive skills are under the control of contingent stimuli. That is, it was an investigation of the effects of praise when that praise was dependent upon the specific content of an answer. The content of the answer was defined as representing a specific cognitive skill. Therefore, if the praise strengthened the target answering behaviour it implied that the underlying cognitive skill was also strengthened.

Organization

This thesis has been organized so that Chapter One reviews literature which provides a foundation for the entire study and gives direction to the work. Chapter Two presents a description

of the initial work which was based on the background theory and literature and a description of the developmental work necessary to produce and test instruments. Chapter Three provides an account of the experimental procedure which developed. Particulars of the procedure for assigning pretests, treatments, and posttests are included in Chapter Three along with a description of the sample and of the data. Chapter Four presents the analysis of data and the specific hypotheses which were investigated. The final chapter, Chapter Five, presents a discussion of the conclusions and implications of the study based on the analysis of the data.

CHAPTER ONE

Background and Related Research

Introduction

This chapter first presents a general discussion of the theoretical background of reinforcement theory as it applies to classroom learning. There follows respectively discussions of: (i) direct reinforcement of classroom behaviour; (ii) vicarious learning; and (iii) factors which influence direct reinforcement and vicarious learning.

Reinforcement and Learning

By definition, oral questioning in classrooms usually involves a question which is asked by the teacher and is answered by a target student. The answer given by the target student may or may not be praised, presumably depending on criteria set by the teacher. Reinforcement theory proposes that the probability that the participating target student will give a similar type of answer in the future is increased if the praise functions as a reinforcer. At the same time, if the exchange between teacher and target is observed by others (as is usually the case, Adams and Biddle, 1970) and the observing student is affected by the praise given to the target, supposedly the probability that the observer will subsequently give an answer that ~~meets~~ meets the criteria for praise is also increased, this time through vicarious reinforcement.

Using both direct and vicarious reinforcement theory terminology, the participating target student who answered the question is called the model and the nonparticipant who watched but did not verbalize is the observer. For the model, the praise following his answer was a stimulus that functioned as a reinforcer provided it increased the frequency of the desired type of answer (Ferster and Perrott, 1968, p. 25). For the observer, the praise following the model's answer functioned as a vicarious reinforcer if it increased the frequency of the

desired type of answer given by the observer. Reinforcement occurs when, as a result of the presentation of a stimulus, the desired behaviour increases. Similarly, vicarious reinforcement occurs when, subsequent to the observation of pleasant consequences for a specified behaviour, the desired behaviour is increased in the observer (Ferster and Perrott, 1968, p. 25). Direct and vicarious reinforcers thus strengthen desired behaviour in the model and observer, respectively. In reinforcement theory, pleasant consequences (reinforcers) are contingent upon the desired behaviour in order to strengthen that behaviour.

It is worth noting that reinforcement as a hypothetical construct is defined by its effects. That is, a given behavioural event occurring in time cannot be identified as a reinforcer until subsequent evidence of a strengthening of the preceding behaviour is obtained. Such a strategy of definition is similar to the linguist's definition of a "yes-no" question whose identity is determined by the answer rather than any characteristic of the question itself. Clearly, there are problems with such an approach to definition. Not only is circularity inherent in the definition but the researcher is also committed to a "post factum" identification of the reinforcing stimulus. The obvious difficulty is to be met herein by using the term P-reinforcer (i.e., potential-reinforcer) when referring to a stimulus that has the potential of functioning as a reinforcer but may eventually prove not to increase the frequency of the desired behaviour. For example, a teacher praises student participation in hopes that the participation will be strengthened. The praise has the potential of strengthening participation although it may not. This praise may be called a P-reinforcer until subsequent evidence substantiates its operation as a reinforcer.

Because of the hypothetical nature of reinforcement, many theories exist to identify the components of the phenomenon. The present study relies mainly on a behaviourist perspective that incorporates aspects from both Skinner (See Skinner, 1953; 1969; 1974) and Bandura (See Bandura, 1969; 1973; Bandura and Walters, 1963) and others with similar inclinations, notably

Ferster and Perrott (1968) and Miller and Dollard (1945). Accordingly Skinner and Bandura will receive greatest prominence in this review of background theory.

Skinner

Skinner describes reinforcement in terms of the relationship between a response (R) and a stimulus (S). If the response is more likely to occur in the future because of the presentation of the stimulus, reinforcement has occurred with the stimulus acting as a reinforcer (S^R); symbolically represented as $R \rightarrow S^R$. He also discusses stimuli in the form of environmental clues to the effect that a pleasant consequence might be available for a particular response, and calls these stimuli discriminative (represented as S^D) to complete the symbolic picture, $S^D \rightarrow R \rightarrow S^R$. For example, a teacher asks a question (S^D) which elicits a specific type of answer (R) that is strengthened by immediate praise (S^R). Additionally with Skinner's description of reinforcement, the amount of strengthening is dependent upon such factors as the schedule of reinforcement (Hilgard and Marquis, 1961; Ferster and Perrott, 1968) and the individual's past history of reinforcement (Kennedy and Willcutt, 1964; Premack, 1965, p. 164). One of the drawbacks of Skinner's notion of reinforcement is that a behaviour (R) must occur before a potential reinforcer can be supplied. That is, he provides an extensive description of the entire process of reinforcement, with reinforcers given directly to participants, and various factors which may interact to influence the amount of learning which occurs through the application of contingent stimuli; but he does not attempt to account for the first occurrence of the desired behaviour except through successive approximations, mutations or neurological explanations. He does not recognize vicarious learning without direct reinforcement of the observed behaviour. He assumes that an observer is engaging in some behaviour, perhaps at a covert level, which has consequences like that of overt behaviour and which may be facilitated through positive reinforcement. According to this view, direct reinforcement can account for all so-called vicarious learning. However,

since Skinner's analysis of these issues (i.e., acquisition of response and vicarious learning) show no clear lines of experimental development at the present time, the theory and research put forth by Bandura proved useful in guiding the present study where vicarious learning is concerned.

Bandura

While there is a large measure of consensus between Bandura and Skinnerian behaviourists over the definition of and function of reinforcers and reinforcement, a major difference is Bandura's discussion of vicarious learning as a mechanism explaining first occurrences of behaviours. Bandura (1969) explains behavioural change through both (i) direct reinforcement and (ii) vicarious learning. A summary of his interpretations of both follows in that order.

Bandura discusses three types of controlling systems for direct reinforcement: (1) Stimulus Control, (2) Internal Symbolic Control, and (3) Outcome Control (Bandura, 1969, p. 45). Stimulus Control, by which behaviour is determined by external stimulus events, is similar to classical conditioning. Internal Symbolic Control postulates a mediational mechanism whereby events are coded and represented in the memory, and behaviour is determined by internal information processing systems. Outcome Control most closely approximates Skinner's view that consequences of behaviour determine the amount of learning. Skinner's view is included, therefore, in Bandura's explanation of direct reinforcement but Bandura also pleads the necessity of additional control systems to provide a more eclectic theory of reinforcement.

Bandura discusses vicarious learning and the influence of reinforcers on vicarious learning. Accordingly, vicarious learning involves (i) symbolic coding and central organization of modelling stimuli, (ii) their representation in memory through "verbal" and "imaginal" codes, and (iii) their transformation from the symbolic forms to motor equivalents (Bandura, 1969, p. 127). It is worth noting in passing that a stricter behaviourist viewpoint would not

recognize these mediational determinants of behaviour.

In Bandura's explanation of vicarious learning, acquisition (i.e. learning) and performance of responses are necessarily distinguished from each other. The acquisition of response depends primarily upon stimulus contiguity and symbolic processes, but the actual strengthening of the performance depends upon the reinforcer given either as a direct reinforcer to the observer, or as a vicarious reinforcer given to the model (Bandura, 1969, p.128). By contrast, Skinner would suggest a neurological answer to the acquisition of response, with direct reinforcement of the observer's behaviour necessary to strengthen the performance.

Bandura (1969) argues that a vicarious reinforcer may act as a discriminative clue by providing information concerning probable consequences under similar circumstances. In a discriminative sense, then, response consequences experienced by another person convey information to the observer about probable contingencies associated with similar behaviours - this parallels with Skinner's description of a discriminative stimulus.

Although the controlling systems may differ, symbolic representation of direct reinforcement according to a Bandurian interpretation can be thought of as basically similar to Skinner's, i.e. $R \rightarrow S$. Bandura's vicarious learning may be represented most simply by: (Observation of) R . In this case, the behaviour of a model serves as a discriminative clue that evokes a similar behaviour in the observer. In addition to this simple case of observational learning, a reinforcer may be given to the model and may act to give added strength to the behaviour of the observer (Bandura, 1969, p. 171, 128). This can be symbolically represented by: (Observation of) $R \rightarrow S^r$. That is, although behaviour can be facilitated by the mere observation of another's performance, reinforcers given to the participant may give additional aid to the facilitation.

More specifically, Bandura's comprehensive theory has four main components all of which influence the nature and

degree of vicarious learning:

1. Attentional Processes - Some individuals are more attentive to the modelling stimuli than others. The complexity of the modelling stimuli and the ability of subjects to discriminate the stimuli affect the attentional processes.
2. Retentional Processes - The symbolic coding operations function here. Interference comes from the rate and organization of the modelling stimuli. Covert practice aids the retentional processes.
3. Motor Reproduction Processes - Motor capabilities must be present for accurate reproduction and limit the amount of reproduction which is possible.
4. Incentive and Motivational Processes - These make individuals attentive to certain cues but not to others and give selective retention.
(Bandura, 1969, pp. 136-143)

Thus, the amount and rate of vicarious learning is dependent upon sensory capacities, motor capacities, and the capacity for representational mediation.

The exact amount of vicarious learning is difficult to predict because observers are not always equally affected by models. For a variety of reasons, some models are more effective than others in any particular situation. These differences in effectiveness can sometimes be attributed to characteristics of both the observer and the model or to the frequency and intensity of reinforcers given to the model (Bandura, 1969). For example, a model who has high prestige has more potential for facilitating learning than one with low prestige. Also, a model similar to the observer will be more efficient in facilitating behaviour than one who is quite different. As for reinforcers, those which are highly desired by the observer will be more efficient as vicarious reinforcers than less desired ones. All of these factors interact to produce a complex situation where vicarious learning may occur but cannot be precisely predicted.

Theory as It Applies to the Present Study

The specific behaviour setting of the classroom is receptive to both Skinnerian and Bandurian approaches. Both can be used to analyse the question and answer dialogue although the processes and learning components differ.

On the one hand, according to Skinner: If a participant receives praise (S^R) from the teacher for some answer (R), the result may be a strengthening of the answering behaviour of the participant, provided praise constitutes an efficient reinforcer. This direct reinforcement is simply $R \rightarrow S^R$. If the observer of this dialogue is subsequently praised for similar types of answers, the result may be a strengthening of the answering behaviour of the observer, again resulting in learning by direct reinforcement. Here, the observed sequence serves merely as discriminative stimulus, and can be represented by: $S^D \rightarrow R \rightarrow S^R$. Vicarious learning without accompanying direct reinforcement of the observer's responses is not recognized in this Skinnerian analysis. The amount of learning by direct reinforcement for both the participant and the observer is dependent upon such factors as how often the praise occurs (the schedule of reinforcement) and whether or not praise has been an effective facilitator previously (the past reinforcement histories of the learners).

On the other hand, according to Bandura, if a participant receives praise from the teacher for a particular type of answer, the result may be a strengthening of the answering behaviour of the participant (direct reinforcement) and a strengthening of similar answering behaviour of the observer (vicarious reinforcement). The amount of the participant's learning is dependent upon the salient stimuli of the environment (Stimulus Control), upon the praise (Outcome Control), and upon the coding of the events (Symbolic Control). The amount of vicarious learning (of the observer) is dependent upon whether or not the observer can sort out what is being praised (Attentional Processes), upon how the observer codes the information gained from the model (Retentional

Processes), upon whether or not the observer is capable of giving the types of answers that gain praise (Motor Reproduction Processes), and upon whether or not the observer is interested in gaining the praise (Incentive and Motivational Processes). All of the control systems and processes may be operating to influence the amount of direct reinforcement and vicarious learning - along with the specific characteristics of the model, observer, and reinforcer.

Direct Reinforcement Studies

Content of Reinforcement Studies

Regardless of whether one takes a Skinnerian or Bandurian stance on reinforcement theory, reinforcers applied directly to participants have been shown to influence a great many kinds of behaviour commonly found in classrooms. These behaviours can be roughly grouped as: (i) those dealing with classroom maintenance, usually related to discipline problems; (ii) those dealing with academic achievement; and, more relevant to the present study, (iii) those dealing with cognitive skill. Each will be dealt with in turn.

On the whole, the studies which are presented hereafter are not exhaustive but serve to illustrate the range of classroom-related behaviour which has been strengthened by direct (and positive) reinforcement, using samples of normal subjects from preschool to university levels. The samples used in the cited studies may appear to give a narrow view of the range of individuals who have been used in reinforcement studies - because in fact, reinforcement techniques have been very helpful to teachers of mentally deficient students, but these will not be discussed in this paper. Additionally, it should be noted that most reinforcement studies deal with contingencies set on an individual basis; therefore, the sample sizes are generally small. However, precise records of the frequency of behaviour, before and after treatment, serve to illustrate the contingent relationships between responses and stimuli to the satisfaction of most researchers in the field.

Bearing in mind the enormous range of reinforcement studies, the studies discussed herein serve only to indicate the variety of classroom-related studies and the need for additional work relating to the strengthening of representations of cognitive skill.

Classroom maintenance research. Disruptive classroom behaviour has been controlled by praising appropriate behaviour while ignoring that which is inappropriate. In an experiment conducted by Allen, Henke, Harris, Baer, and Reynolds (1970), hyperactivity in a four-year-old boy was controlled by praising successively longer attention spans. The result was an increase in attention span and a complementary decrease in hyperactivity.

In a similar vein, tardiness was reduced in a fifth-grade classroom by giving a reinforcer to students who were in the classroom on time. The reinforcer consisted of posting the punctual students' names on a chart (Hall, Cristler, Cranston, and Tucker, 1970). Although not investigated specifically, this study implied the presence of vicarious reinforcement, since the names on the chart could reinforce both directly and vicariously.

In another study peripherally related to classroom maintenance, a five-year-old girl increased interactions when praise was contingent upon cooperative play (Hart, Reynolds, Baer, Brawley, and Harris, 1971). Similarly, socially acceptable responses to peers, such as working cooperatively, were strengthened in a four-year-old boy, while unacceptable behaviours, such as hitting, were decreased by means of reinforcing the acceptable responses with praise and ignoring unacceptable responses (Scott, Burton, and Yarrow, 1969).

Academic achievement research. Besides classroom maintenance studies, academic performances have been the target of some reinforcement studies. For example, spelling, mathematics, social science, and science skills were strengthened in a study by Evans and Oswalt (1971). In four different experiments they demonstrated that if early dismissal from class or story reading

depended on their performance:

1. Fourth-grade students spelled words correctly more often.
2. Fourth-grade students correctly completed mathematics problems more often.
3. A sixth-grade student correctly answered questions about a social science lesson more often.
4. A sixth-grade student correctly answered questions about a science lesson more often.

Adding support to the previous study, Zimmerman and Zimmerman (1966) showed that an 11-year-old boy of normal intelligence spelled words correctly more often when the attention of the teacher (the reinforcer) depended on the correct spelling. Yet another study dealing with the reinforcement of academic performances demonstrated that the addition skills of fourth and sixth-grade students improved if praise depended upon the improvement (Hurlock, 1925). And, Staats and Butterfield (1965) demonstrated that a 14-year-old boy increased his oral and silent reading skills as well as his vocabulary skills when reinforcers were given for specified behaviour; while Hart and Risley (1968) have used reinforcers to influence the language skills of disadvantaged preschool children by making access to play materials contingent upon a specific type of adjective-noun combination.

Attitudes, somewhat related to both classroom maintenance and academic performance, were strengthened in a study by Bostrom, Vlandis, and Rosenbaum (1961). University students, who had written essays defending positions on attitude-related issues that were contrary to their previously assessed positions, were randomly assigned good marks and poor marks. Evaluation showed that verbal reports of attitudes were reinforced by good marks but not by poor marks.

Cognitive skill research. Apart from studies dealing with classroom maintenance and academic performances, relatively few studies have been directly concerned with the strengthening of

cognitive skill in its variety of manifestations.¹ (See MacMillan, 1973, p. 210 for further comment). However, one such study undertaken by Goetz and Baer (1973) demonstrated that diversity in block building, which could be thought of as a form of diversity in thinking, was strengthened by contingent praise. In their study, three preschool girls were provided with descriptive praise, e.g., "Oh, that's very nice - that's different", for new patterns of construction made with blocks. Results showed that with contingent praise the subjects produced new patterns at a higher rate than before reinforcement or when they were praised for repeating old patterns.

Although current research is heavily slanted toward discipline-related behaviour, the preceding studies have a common theme - the strengthening of desired classroom behaviour through the application of a direct reinforcer. They showed that maintenance behaviours, academic achievement, and cognitive skill can all be affected by direct reinforcement and imply that a greater variety of behaviours are controlled by contingent reinforcers.

Vicarious Learning Studies

Content of Vicarious Learning Studies

Although direct reinforcers are capable of strengthening a multitude of behaviours, when the operant level of a behaviour is very low, direct reinforcement is usually ineffectual. However, the same behaviours can be manifested through vicarious learning. It has been demonstrated that many classroom-related behaviours with relatively low operant levels have been acquired in this way. Examples range from moral judgments to those more related to academic performance and cognitive skill; they will be discussed in that order.

1 Staats (1968) maintains that academic performance is closely allied with cognitive skill when illustrating the acquisition of writing, reading and number concept in his young daughter.

Moral judgment research. Moral judgment responses were the focus of a vicarious learning study by Bandura and McDonald (1963). In it, some of the children observed adults who expressed moral judgments counter to the group's orientation while others, not watching an adult, were given approval directly when the desired judgments occurred. Children in one sub-sample received approval for adopting the adult's judgments (vicarious learning and direct reinforcement) while others in another sub-sample watched the adult but were not given approval (vicarious learning only). The results showed that vicarious learning, both with and without direct reinforcement, produced the desired types of judgments more often than approval given directly for the desired moral judgments.

Academic achievement research. Specific language paradigms constituted the behaviour desired in a vicarious learning study by Rosenthal and Whitebook (1970). Third and fourth-grade students initially observed a live model who verbalized specific sentence patterns, words, and verb tenses in response to a set of pictures. Subsequently the students showed a significant increase in the types of verbalizations similar to those given by the model even though they had received no P-reinforcer. Although this research seems somewhat inconsequential on the surface, it does serve to illustrate vicarious learning of language skills in a controlled situation where no overt P-reinforcers were offered either to the model or to the observer.

Cognitive skill research. Somewhat more related to the goals of the present study are studies illustrating question-asking learned vicariously. Question-asking is regarded as a significant cognitive skill (See Suchman, 1964; Cazden, 1970) in the same way that question-answering is presently. Question-asking was the target behaviour in a study by Rosenthal, Zimmerman, and Durning (1970) in which they demonstrated that sixth-grade children increased questions about physical attributes, functional uses, causal relationships, and value judgments by watching an adult model ask these types of questions. In this case, no praise or other P-reinforcer was given to the models or subjects. The frequency of the desired types of

questions increased for the observers presumably because of their observation of the model's behaviour. In a similar study, Henderson and Garcia (1973) increased question-asking in six-year-olds by allowing them to watch a model who asked questions. A confounding factor, however, is that subjects were instructed to ask questions rather than being given the opportunity to manifest the behaviour spontaneously.

Question-asking was also the focus of attention for Krumboltz and Thoresen (1964). In their study, one sample of 11th-grade students, having watched a videotape of a model, was then praised for asking questions similar to the model's. Another sample did not watch the videotape but was praised for manifestations of the appropriate question-asking behaviour. A control group received no treatment. The results showed: (i) all subjects who were praised, regardless of whether or not they watched a model, displayed more of the desired behaviour than the control group; (ii) male subjects in the videotape treatment group subsequently displayed more of the desired question-asking than males in the "praise only" group. Similar results also occurred in a later study by Krumboltz and Schroeder (1965).

The preceding studies illustrate the effectiveness of vicarious learning in bringing about behaviours which would otherwise have a low operant level. They show that the mere observation of a model's behaviour can be sufficient to induce the desired behaviour (Henderson and Garcia, 1973; Rosenthal and Whitebook, 1970; Rosenthal, Zimmerman, and Durning, 1970). Moreover, approval given for behaviour similar to the model's, aids in strengthening the behaviour (Bandura and McDonald, 1963; Krumboltz and Schroeder, 1963; Krumboltz and Thoresen, 1964). Taken together, the research implies that a greater variety of behaviours than those specifically investigated, including those related to cognitive skill, are and can be learned vicariously. In fact, a vicarious stimulus is sometimes more efficient than the application of direct P-reinforcers at strengthening a low frequency behaviour.

Factors Influencing Direct Reinforcement
and Vicarious Learning

Due to the number of variables involved in the processes of both direct reinforcement and vicarious learning, there are a multitude of factors which may influence the amount of learning which occurs. A sample of these influential factors relating to direct reinforcement and vicarious learning will be presented in that order. The factors are not exhaustive, but serve to illustrate the variety of influential stimuli.

Direct Reinforcement

In cases of direct reinforcement, i.e. $R \rightarrow S^R$, characteristics of the stimuli, the response, and the reinforcer, as well as learner and environmental characteristics all interact to influence learning. For example, research has shown the following to be influential:

(i) specific characteristics of the reinforcing stimuli, such as percentage, intermittance, and magnitude of reinforcement (See for example, Bandura, 1969, p. 30; Ferster and Perrott, 1968) and different kinds of reinforcers (See for example Travers, Reid, and Van Wagenen, 1964); (ii) specific characteristics of the subjects themselves, such as sex of the subject (See for example Trent and Cohen, 1973) and past history of reinforcement (See for example Kennedy and Willcutt, 1964); and (iii) awareness of contingencies on the part of subjects (See for example Dulany, 1961).

Additional research is needed to explore thoroughly the salient features of the system, such as characteristics of those who provide the P-reinforcers (See McDonald, 1973) and to discern whether awareness is truly necessary for conditioning (See Bandura, 1969, pp. 568-557). However, it seems in general that efficient reinforcing stimuli are available for most individuals in many different situations, and these reinforcing stimuli are capable of strengthening a multitude of behaviours, regardless of and in spite of specific characteristics of the feedback agent and of the desired response.

Vicarious Learning

Since an observer may be reinforced directly for performing acts similar to those of a model, all of the influential factors affecting direct reinforcement may be available to affect the amount of learning (Bandura, 1969). Additionally, specific characteristics of the model and observer may also affect the amount of learning which occurs vicariously, as may characteristics of the reinforcing stimuli given to the model.

Observer characteristics. Observer characteristics that have been shown to be influential include: (i) dependency (Jakubczak and Walters, 1959; Kagan and Mussen, 1956; Ross, 1964); (ii) self-esteem (de Charms and Rosenbaum, 1960; Gelfand, 1962; Lesser and Abelson, 1959); (iii) level of competence (Kanareff and Lanzetta, 1960b); (iv) sex (Bandura, Ross, and Ross, 1969; Maccoby and Wilson, 1957; Krumboltz and Thoresen, 1964; Rosenbaum and Bruning, 1969); and (v) past experience of being rewarded for matching behaviour (Miller and Dollard, 1945; Schein, 1954).

Model characteristics. Attributes of the model that have been shown to be influential include: (i) high competence (Gelfand, 1962; Mausner, 1954a, b; Mausner and Bloch, 1957; Rosenbaum and Tucker, 1962); (ii) expertise (Mausner, 1953); (iii) age (Bandura, Ross, and Ross, 1963a; Maccoby and Wilson, 1957; Rosenblith, 1959, 1961); (iv) status (Lefkowitz, Blake, and Mauton, 1955); (v) social power (Bandura, Ross, and Ross, 1963b, Mischel and Grusec, 1966); (vi) ethnic status (Epstein, 1966); and (vii) attractiveness (Bandura and Huston, 1961; Grusec and Mischel, 1966).

Vicarious reinforcers. Directly related to the present study, vicarious reinforcers have been shown to affect a great variety of behaviour commonly found in classrooms and a number of these are discussed below. Once again, the studies are grouped according to their concern with (i) classroom maintenance, (ii) academic achievement, and (iii) cognitive skill.

Classroom maintenance research. Children have been shown to increase aggressive behaviour after watching another individual receiving P-reinforcers for aggressive behaviours (Bandura, 1965;

Bandura, Ross, and Ross, 1969). Although these experiments were not carried out in a regular classroom environment, they illustrate how such aggressive behaviours may be learned in the classroom. To illustrate, in the Bandura (1965) study, nursery school children watched a videotape of a model who (after punching, hitting, and kicking a big doll while making aggressive verbalizations toward it) received praise, candy, and drink. Subsequently when given the opportunity, children who had observed pleasing consequences of the model's behaviour performed similar aggressive behaviours more often than those who had observed a model punished for similar aggressive acts.

Attentiveness, or steady application to work, also related to classroom maintenance, has been influenced by vicarious reinforcers. Broden, Bruce, Mitchell, Carter, and Hall (1970) showed that while a teacher was systematically giving her attention (direct reinforcer) to one student for appropriate attending, the attending behaviour of an observing student increased too (vicarious reinforcement).

Another classroom maintenance behaviour which has been influenced by vicarious reinforcement is social behaviour. O'Conner (1969) showed that preschool isolates increased their social interactions after watching a film depicting children's social interactions that resulted in positive consequences - the positive consequences functioned as vicarious reinforcers.

Academic achievement research. Academic skills have also been strengthened through the use of vicarious reinforcers. Travers and Van Wagenen (1964) showed the effect of vicarious reinforcers on the learning of German vocabulary. Here the extent of learning for observers was dependent upon the amount of praise given to the participants, that is, the number of correct responses. In subsequent trials, observers were given a test similar to the model's but employing different English words. The results showed that the observers gave more correct German words than might have been expected by chance alone.

Further support to the effectiveness of vicarious reinforcement of academic behaviours is shown in studies by

Ditrichs, Simon, and Greene (1967); Kanfer and Marston (1963); Marston (1964); Marston and Kanfer (1963); and Smith and Marston (1965). Each study dealt with vicarious reinforcement of the use of a specific class of words. In general, the results showed that the use of a specific class of words can be increased in an observer by his watching another being praised for use of them.

Cognitive skill research. As an example of cognitive skill being influenced by vicarious reinforcement, Krunboltz, Varenhorst, and Thoresen (1967) demonstrated the strengthening of information-seeking in high school girls. The girls watched a videotape of another student being praised for seeking information. Subsequently the observing students displayed more of the desired behaviour than did a control group.

The preceding studies have reported on a variety of classroom related studies that have illustrated vicarious reinforcement. They yield evidence that an observer's behaviour can be strengthened by watching the consequences of a model's behaviour. Additionally, the amount of influence of a vicarious reinforcer depends upon the rate, number, distribution, and complexity of the modelling stimuli and the discriminability of the modelling stimuli. Taken together, the research illustrates the generality of vicarious reinforcement as a technique for strengthening classroom behaviours but also indicates many of the complexities and variables which interact to influence the amount of learning.

Summary of Influential Factors

The components of direct reinforcement, vicarious learning, and vicarious reinforcement somewhat determine the variables which may affect the amount of learning in each case. For each, specific characteristics of the subjects, the stimulus, and the response have the potential of governing the amount of learning. Additionally, for direct and vicarious reinforcement, the reinforcing stimulus has specific characteristics, such as frequency and magnitude, which may govern the amount of learning. Awareness of the contingencies of reinforcement (the dependent relationship between the response and the reinforcing stimulus) may also affect the potency of the

reinforcer. In vicarious situations, the particular aspects of the model may be influential. Because of the number of components involved, alterations of any of them may demonstrate differences in the amount of learning.

Summary of Research Dealing with
Direct and Vicarious Reinforcement

Reinforcement theory as it would apply to classroom answering behaviours implies that a selected type of answer from a participant can be strengthened by an appropriate reinforcing stimulus. The selected type of answer, this time given by an observer of the direct reinforcement, could also be strengthened by vicarious reinforcement. Evidence to this effect is summarized in Table 1. Relevant studies are classified alphabetically as (i) those demonstrating direct reinforcement, and (ii) those demonstrating vicarious learning. Within each of these broad categories, a brief description of the behaviour which was strengthened is supplied along with a brief description of the reinforcer used and the type of sample. The range of behaviours strengthened through direct reinforcement varies from cooperative play to diversity in block building. Similarly with vicarious learning, strengthened behaviours vary from moral judgments to asking questions. The reinforcers, both direct and vicarious, include such things as food and praise, with a great majority using praise. And finally, the samples include normal individuals over a great range of ages. In general, the summarized research shows that under experimental control, direct and vicarious reinforcers have strengthened classroom maintenance behaviours, academic achievement, and cognitive skill using many different types of reinforcers with a large range of individuals.

TABLE 1

Summary of Research Dealing with Direct Reinforcement and Vicarious Learning

| | Pupil Behaviour | Direct Reinforcer | Vicarious Reinforcer | Sample | Investigators |
|-------------------------|--|----------------------|----------------------|------------------------------|------------------------------|
| DIRECT REINFORCEMENT | Hyperactivity | Praise | | 4-yr old boy | Allen et al. (1970) |
| | Attitudes | Good marks | | Univ. stud. | Bostrom et al. (1961) |
| | Spelling, Math. Soc.Sci. & Sci. skills | Early dis. & stories | | 4th & 6th gr. students | Evans & Oswalt (1971) |
| | Diversity in block build. | Praise | | Preschool girls | Goetz & Baer (1973) |
| | Prompt to class | Names on chart | | 5th grade students | Hall et al. (1970) |
| | Co-op play | Praise | | 5-yr old boy | Hart et al. (1971) |
| | Language skills | Play mat. | | Disad. preschool. | Hart & Risley (1968) |
| | Add. skills | Praise | | 4th & 6th gr. students | Hurlock (1925) |
| | Soc.accept. behaviour | Praise | | 4-yr old boy | Scott et al.(1969) |
| | Reading, writ. & math. | Toys and praise | | 1-4 year old girl | Staats (1968) |
| | Reading & vocab. skills | Tokens | | 14-year old boy | Staats & Butterfield (1965) |
| | Spelling skills | Attention | | 11-year old boy | Zimmerman & Zimmerman (1962) |
| | Aggres. behaviour | | Praise & food | Nursery sch. children | Bandura (1965) |
| | Moral Judge. | Praise | | Children | Bandura & McDonald (1963) |
| | VICARIOUS LEARNING | Aggres. behaviour | Incentives | Toys & food | Nursery sch. children |
| Attending to work | | Praise | Praise | School child. | Broden et al. (1970) |
| Hostile verbs | | | Praise | 7,8, & 9th gr. students | Ditrichs et al.(1967) |
| Question asking | | | | 6-yr old child | Henderson & Garcia (1973) |
| Specific class of words | | | Praise | University students | Kanfer & Marston (1963) |
| Question asking | | Praise | Praise | 11th gr. students | Krumboltz & Schroeder (1965) |
| Question asking | | Praise | Praise | 11th gr. students | Krumboltz & Thoresen (1964) |
| Info. seeking | | | Praise | Hi.School girls | Krumboltz et al. (1967) |
| Human nouns | | | Praise | University students | Marston (1964) |
| Specific class of words | | | Praise | Male Univ. students | Marston & Kanfer (1963) |
| Soc. interact. | | | Pos.Conseq. | Nursery sch. children | O'Conner (1969) |
| Fast lever movements | | Candy | Candy | 5-7 yr. olds | Rosenbaum & Bruning (1969) |
| Sent. patterns | | | | 3rd & 4th gr.stud. | Rosenthal & Whitebook (1970) |
| Question asking | | | | 6th gr.stud. | Rosenthal et al.(1970) |
| Specific class of words | | | Praise | Male Univ. students | Smith & Marston (1965) |
| German vocabulary | | Praise | 4,5,6th gr. students | Travers & Van Wagenen (1964) | |

Discussion

The research reported here generally upholds the theories of direct and vicarious reinforcement and vicarious learning. Both the research and theory imply that many variables interact to influence the amount of learning which occurs in direct and vicarious learning situations. Even with confounding features like those of the studies dealing with the sex-related behaviour of aggression (Bandura, Ross, and Ross, 1969) it is possible to create the right contingencies to result in learning - the female subjects who had previously been trained not to be aggressive in our culture, learned aggression vicariously. Although many of the influential variables, such as sex of the subjects and the specific schedules of reinforcement, have been subjected to control in some research, a multitude of potentially influential components, such as characteristics of those who give reinforcers, have been left unexamined (McDonald, 1973). Many, such as awareness of contingencies of reinforcement, are in need of further exploration (Bandura, 1969, pp. 568-577). Nevertheless, based on the research, it seems that contingencies can often be set to ensure that most individuals learn the desired task and, in spite of the multitude of potentially influential variables, suitable contingencies of reinforcement can be devised to strengthen the desired behaviour principally by identifying preferred activities or items which may act as reinforcers and by specifying the circumstances under which these P-reinforcers may be given.

The research discussed herein has indicated that direct and vicarious reinforcement are potentially useful teaching techniques in that they have been shown to strengthen a variety of behaviours, in a variety of individuals, under a variety of contingencies. Most of the studies cited, however, deal with maintenance behaviour rather than with academic achievement or cognitive skills. Although opinion differs as to the explicit functions of education (See Peters, 1966, p. 27; 1965, pp. 97-100; 1967, pp. 6-7), it seems that one of those functions is the

development of cognitive skills. Accepting this, the influence of P-reinforcers on manifestations of cognitive skill is as important a realm of investigation as is the influence of P-reinforcers on the more maintenance-related behaviour. More research in this line is needed to complement that already existing in order to test the reinforcement framework as an explanation for classroom behaviour.

CHAPTER TWO

Procedure - Design Development

Introduction

This chapter describes the research problem as it was first formulated, the procedures initially undertaken, the modifications that were required, and the circumstances that necessitated them.

Statement of the Problem

In the beginning, the major problem was whether cognitive skill, in the form of specifically selected answering behaviour, could be shown to be affected by direct and vicarious stimuli which had the potential of being reinforcers. Justification of the study thus formulated was two-fold. First, cognitive skill as a realm of investigation deserved attention since teachers are continually plagued with problems of how best to teach specific cognitive skills as one of the main emphases of classroom objectives (Bloom, 1956, p.7). Second, cognitive skill as a realm of investigation deserved attention since although psychologists have demonstrated direct and vicarious reinforcement of a multitude of behaviours the realm of cognitive skills has remained largely neglected.

The elusive concept of cognitive skill can be defined as: a specific type of intellectual ability (Bloom, 1956, p.7) or competence (Henderson and Garcia, 1973). Although such a definition does not avoid the problem of how to measure the covert intelligent behaviour, behavioural manifestations can be postulated as overt representations of specific intellectual abilities and competencies and, therefore, of a specific cognitive skill (Skinner 1974). More specifically, if an example of cognitive skill can be identified by specific answering behaviour, and if reinforcement theory holds true for such behaviour, it should be possible to show the strengthening of the answering behaviour and of the particular cognitive skill represented.

On the whole, studies have shown the reinforcement framework to be a satisfactory explanation of the strengthening of the behaviours they investigated. However, considering the whole range of reinforcement studies, relatively few, with the exception of studies such as those by Henderson and Garcia (1973) and Goetz and Baer (1973) who investigated question-asking and diversity in block

building, respectively, have investigated the effect on cognitive skill. Although question-asking and diversity in block building may be valid manifestations of cognitive skill, they do not consume as much student time as does answering questions (Adams and Biddle, 1970). It seems that a complete look at the influence of reinforcers on cognitive skill ought to include an exploration of the effects of reinforcers on answering behaviours.

For the purposes of the present investigation, specific types of answers to questions were assumed to be manifestations of a student's cognitive skill. It followed then that from an examination of a student's answers it was possible to infer specific characteristics of his cognitive skill. Consequently, it was also possible to arrange contingencies so that some types of answers received Potential-reinforcers (P-reinforcers) and others did not. In this manner, it was conceivable to test the existence of direct and vicarious reinforcement of manifestations of cognitive skill through an investigation of the effects on selected answering behaviours.

Bearing in mind that some psychologists (See Hilgard and Marquis, 1961) discuss at length those cognitive structures which determine intelligence, it should be noted that the present interpretation of cognitive structures is that such structures can be inferred only so much as common physical representations of the structures are present to examine. That is, cognitive skill is private and cannot be discussed unless common verbalized descriptions are agreed upon to illustrate some specific cognitive skill. It is possible to measure such covert behaviours only through accompanying overt analogues (See Skinner, 1969). It is not adequate or necessary to talk of cognitive structures without discussing overt manifestations of behavioural representations of them through which consensus would define those structures.

From the onset of the present study, the desire to concentrate on cognitive skills manifested in answering behaviour brought about several problems in developing the design of the study. Initially, the basic plan was to arrange a situation wherein answers could be

solicited from a participating subject and P-reinforcers could be given for specific types of answers, while the whole procedure was viewed by an observer. In this manner the existence of direct and vicarious reinforcement could be tested. The necessary prerequisites for proceeding were the devising of methods for (i) classifying various answers as manifestations of specific types of cognitive skills, and (ii) devising methods for soliciting specific types of answers.

Classifying Questions and Answers

Whether or not children, or adults for that matter, have achieved cognitive competency can only be gauged from their manifested behaviour. In classrooms most of that manifested behaviour is verbal and most occurs in answer to questions or demands. It follows then that accurate judgment about cognitive competency can only follow from the identification of "cues" in those verbal responses. Necessarily, such cues must be differentiated and defined so that much of the extent and variety of cognitive competency is encompassed. There was need then to classify answers as to the specific cognitive skills they represented. Once answers could be identified as representing a specific cognitive skill, instruments could, theoretically, be developed which would measure the cognitive skill of an individual. Initial concern then was with arriving at an adequate cognitive skill classification scheme and at developing instruments adequate for measuring that cognitive skill.

Bloom's Taxonomy of Educational Objectives.

The Taxonomy of Educational Objectives (Bloom, 1956) comprises a classification of cognitive skills which ostensibly ought to meet the needs of the present study and which initially appeared ideal for use. The Taxonomy was designed as a guide for the analysis of educational outcomes in the cognitive areas of remembering, thinking, and problem solving (Bloom, 1956, p. 2) and was formulated by a committee which systematically surveyed educational objectives and analyzed those objectives according to their intended outcomes.

The committee also attempted to arrange the objectives in a hierarchical order - from simple level of thinking to complex. Categories of the Taxonomy and brief summaries of each category follow:

1. Knowledge-Ability to remember previously learned material.
 2. Comprehension-Ability to grasp the meaning of material.
 3. Application-Ability to use learned material in new and concrete situations.
 4. Analysis-Ability to break down material into its component parts so that its organizational structure may be understood.
 5. Synthesis-Ability to put parts together to form a new whole.
 6. Evaluation-Ability to judge the value of material for a given purpose.
- (Gronlund, 1970, p. 20)

The hierarchical nature of the Taxonomy presupposes that higher levels of objectives build on the adjacent lower level, and that mastery of an objective at one level implies mastery at all lower levels.

Although the developers of the Taxonomy acknowledge the fact that the categories of objectives are somewhat arbitrary and many are not well differentiated from each other (Bloom, 1956, pp. 15-16, 24), the Taxonomy is seen by some as a useful scheme for classifying classroom objectives (See Bloom, Hastings, and Madaus, 1971, p. 40; Gronlund, 1970, pp. 18, 24; Kibler, Barker, and Miles, 1970, p. 45; Tanner, 1972, pp. 7-20). While classifying objectives was the prime purpose of the Taxonomy it should none the less be possible to classify both questions and answers according to which level of objective they represent (See Bloom, Hastings, and Madaus, 1971, p. 40). That is, questions representing a specific level of objective can be classified according to the specific cognitive level, and correct answers to the questions can also be classified as belonging to the specific cognitive level. Therefore, working from the Taxonomy, not only should classroom objectives be susceptible to classification as to their level of cognitive skill, but classroom questions and answers should also be susceptible to

such classification.

Research using the Taxonomy. The Taxonomy has appeared in many studies as a convenient and useful tool. For example, Bartolome (1969) used it to compare teachers' objectives and the questions teachers employed, finding incongruity between the cognitive level of objectives set down and the supposed level of questions they used. The Taxonomy was also used by Hunecke (1970) in finding that teachers using and not using curriculum guides, did not differ significantly in the level of objectives and questions they employed.

Hunkin's (1969) study was more specific. He investigated whether a dominant use (in social studies materials) of Analysis and Evaluation level questions as opposed to Knowledge level questions, would improve the cognitive level of achievement. The students who used the Analysis level and Evaluation level materials did better in the area of Evaluation on a posttest than those who used Knowledge level materials, but no differences were found to exist between the groups for the Knowledge, Comprehension, Application, Analysis, and Synthesis areas. In another similar study, this time exploring the effects on critical thinking, Hunkins (1970) found, contrary to the previous results, that students who used social studies materials that emphasized Analysis and Evaluation level questions did no better on a posttest measuring critical thinking than those who used material emphasizing Knowledge level questions.

Apparently, in the above studies, the Taxonomy could be used satisfactorily to classify the cognitive level of objectives and of questions which corresponded to specific objectives - at least to the satisfaction of the researchers themselves. The implication was that the Taxonomy was a satisfactory device for such a purpose. On the strength of such precedents and the theoretical elegance of the Taxonomy, it seemed reasonable to accept the dimensionalization of the cognitive domain as presented by the Taxonomy.

However, if it is true that any inference about the cognitive skill of an individual can only be drawn from some of his manifested behaviour, then any classificatory scheme for describing the cognitive

condition of individuals must have its corresponding behavioural analogue to be measurable. Therefore, the immediate problem for the study was to arrive at a set of behavioural specifications that might presumably be seen as corresponding with the categories specified in the Taxonomy. In other words, the next task was to develop questions which represented objectives at all levels of the Taxonomy so that a correct answer could be recognized as representing a specific cognitive level.

Questions Based on the Taxonomy. To develop appropriate questions, a variety were designed to represent all levels of the Taxonomy. The content of the questions was derived from a single stimulus device. The film, Reflections on Time (Encyclopaedia Britannica Corp., 1970), was used because it provided a range and variety of visual and verbal stimuli that in themselves covered the gamut of Bloom's cognitive behaviours and allowed for exact reproduction of the stimulus material for all subjects. The film, designed to accompany a high school earth science lesson, was considered to be of general interest and generally unfamiliar to all subjects.

The basic forms of questions used in classrooms include short-answer completion items, multiple-choice items, and oral completion items (Hunkins, 1972). Each of these forms of questions theoretically has the potential of soliciting different levels of answers according to the Taxonomy, and the relative merits of each needed to be explored. The first form to be explored was the short-answer completion question.

Short-answer completion items. Items, based on the film and representing specific objectives defined by the Taxonomy, were designed for each cognitive level. By using behavioural terms and forms such as those suggested by Gronlund (1970), Hunkins (1972), and Sanders (1966), short-answer completion items were constructed. Table 2 presents examples of the behavioural terms for each level of objective as set out by Gronlund (1970, pp.20-21) that were used as guidelines.

TABLE 2

Behavioural Terms for Each Level of
the Taxonomy

| Level of Objective | Behavioural Terms |
|--------------------|--|
| 1. Knowledge | Define, describe, identify, label list, match, name, outline, reproduce, select, state |
| 2. Comprehension | Convert, defend, distinguish, estimate, explain, extend, generalize, give example, infer, paraphrase, predict rewrite, summarise |
| 3. Application | Change, compute, demonstrate, discover manipulate, modify, operate, predict, prepare, produce, relate, show, solve, use |
| 4. Analysis | Break down, diagram, differentiate, discriminate, distinguish, identify, illustrate, infer, outline, point out, relate, select, separate, subdivide |
| 5. Synthesis | Categorize, combine, compile, compose create, devise, design, explain, generate, modify, organize, plan, arrange, reconstruct, relate, reorganize, revise, rewrite, summarize, tell, write |
| 6. Evaluation | Appraise, compare, conclude, contrast, criticize, describe, discriminate, explain, justify, interpret, relate summarize, support |

Examples of the 40 questions which were designed according to Gronlund's criteria to correspond to each of the categories of the Taxonomy are found in Table 3.

TABLE 3

Examples of Short-Answer Completion Items
Arranged Hierarchically According to the
Taxonomy of Educational Objectives

| Level | Example Item |
|---------------|--|
| Knowledge | Which is more accurate, a cesium clock or a sundial? |
| Comprehension | Paraphrase the expression "The clock of time is wound but once." |
| Application | Draw a graph to show man's place in the world of time. |
| Analysis | Identify scenes in the film that illustrate the unknown qualities of time. |
| Synthesis | Define time in your own words. |
| Evaluation | Evaluate the statement that "Time is as constant or variable as one makes it." |

The remaining items used in this initial work are found in Appendix A.

Procedure. To test the efficacy of short-answer completion questions in soliciting different levels of answers and measuring cognitive skill, eight volunteer subjects (aged 14 to 31 years) were used. Five of them watched the entire film and three watched only Part II and III of the film. They were then required to write answers for six to eighteen of the short-answer questions. The specific questions completed by each subject are found in Appendix A.

Questions for each subject varied as to: (i) content (in order to test as many items as possible); (ii) arrangement (in order to determine if random or hierarchical arrangement was preferable); (iii) number (in order to determine a reasonable length for a measuring device); and (iv) parts of the film covered (in order to determine if a sufficient number of questions could be designed about only Parts II and III of the film thus decreasing the viewing time).

Results. After each subject completed the items, he was interviewed by the investigator. In the interview, subjects were asked to evaluate critically the questions they answered as to wording and arrangement of items. On the basis of comments some items were eliminated while others were reworded for further use. The three subjects who had items arranged randomly according to the hierarchy criticized the arrangement of items by stating that items requiring longer and more complicated answers (which were usually the higher level items) more logically ought to be presented at the end of the list while those items requiring shorter answers (usually the lower level items) would be better placed at the beginning. Additionally, completion time for the subjects varied from four minutes for six items to 25 minutes for 18 items.

This initial work provided a number of suitable short-answer completion items for use in the next stage of the operation, and indicated subjects' preferences for a hierarchical arrangement of items as opposed to a random arrangement.

Revised short-answer completion items. Following revision of a number of items, the elimination of others, and the addition of some more, a new instrument to test the cognitive level of subjects was constructed. The number of items was increased to 24 items since, in general, a longer instrument makes a more reliable one.

Although the previous work indicated a preference for hierarchically arranged items, it was thought that with the increased length of the instrument, the order of presentation of items might influence the way in which later items were answered. If subjects were fatigued by the time they reached the later items, the measure of their cognitive skill at the higher levels would not be accurate. This would also interact with and be influenced by a seriation effect (See Hovland, 1938; Jensen, 1962; McGeoch and Irion, 1952, p. 155; Mednick, 1964, pp. 57-60) wherein items at the beginning and end of a list are often recalled and completed more readily than those in the middle.

The cognitive levels of objectives as represented by the items were randomized in six different arrangements, thereby making it possible to investigate the effect of the order of presentation of items regardless of whether the effect was due to fatigue or seriation. The six different forms of this revised Short-Answer Test are found in Appendix A.

Procedure and results. Each form of the Short-Answer Test was assigned to one university student, after the student had watched Part II of Reflections on Time. If the Taxonomy was a true hierarchy and the questions validly represented each level of the Taxonomy, those who had high scores at a particular level should also have high scores at all lower levels. Table 4 presents the number of items that were correctly completed at each level of the Taxonomy.

TABLE 4
 Number of Correct Completion Answers at Each
 of the Six Levels of the Taxonomy with a
 Maximum Score of 4 at Each Level

| Subject | Level | | | | | |
|---------|-------|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | 4 | 2 | 4 | 3 | 2 | 4 |
| 2 | 1 | 2 | 3 | 3 | 2 | 2 |
| 3 | 1 | 1 | 2 | 1 | 3 | 3 |
| 4 | 3 | 1 | 2 | 4 | 3 | 3 |
| 5 | 3 | 3 | 2 | 3 | 1 | 4 |
| 6 | 4 | 1 | 2 | 3 | 3 | 2 |

The results show that those who answered all items correctly for objectives of a specific level did not necessarily answer correctly all of the items at subsuned lower levels.

Discussion. It was impossible to label a subject as operating at a specific cognitive level based on the test data. The results indicated that the revised Short-Answer Test did not provide a satisfactory measure of cognitive skill. Additionally, the increased length of the revised test did not add to its effectiveness as a measuring instrument. That is, it was no more possible to identify an optimum level of operation using 24 items than it was using the preceding shorter tests.

The order of presentation of items in the test did appear to affect the answers in that shorter answers were usually given to questions at the beginning of the test. Additionally, most subjects were hurried to complete the 24 items in the 25 minutes allotted. Many forms of the test with random arrangements of items would control for the influence of the order of presentation of the items and an instrument shorter than 24 items was recommended.

Although the results shed doubt on the hierarchical nature of the Taxonomy, the form of items used may have contributed to some of the confusing results. That is, using short-answer completion items, it was sometimes difficult to determine if an answer was correct and adequate because of the great variety of answers which could be given for any one item. The subjectivity of judging the correctness of answers at each level had also occurred in the preceding work only to be supported by the analysis of answers to the revised Short-Answer test. To rectify this problem of subjectivity of scoring in an attempt to obtain a more objectively scored test, additional work was required using different item forms.

Multiple-Choice items. Although the short-answer completion items of the preceding work proved conditionally satisfactory in that they elicited answers over the entire spectrum of the Taxonomy, the subjectivity of the scoring was undesirable. Multiple-choice items theoretically have a potential for permitting the objective scoring which was needed.

The multiple-choice items were constructed by using behavioural terms such as those suggested by Gronlund (1970), Hunkins (1972), and Sanders (1966) and by employing some of the previously used short-answer items. A test, based on Part II of Reflections on Time, was designed in which the first four levels of the Taxonomy were represented by 4-choice items. Since Synthesis and Evaluation levels require a unique creation not possible within the multiple-choice framework (See Sanders, 1966, p. 126), it is virtually impossible to test them with multiple-choice items. Therefore, the last two levels were represented by short-answer completion items taken from preceding work. The items were arranged hierarchically according to the Taxonomy. In Table 5 examples of the multiple-choice items for the first four levels of the Taxonomy are listed. In order to test as many items as possible, two forms of the Multiple-Choice Test were devised. They are found in Appendix A.

TABLE 5

Examples of Multiple-Choice Items Arranged
According to their Cognitive Level of the
Taxonomy of Educational Objectives

| Level | Item |
|------------------|---|
| 1. Knowledge | <p>Which of the following best describes Stonehenge?</p> <ul style="list-style-type: none"> a. An early calendar b. An early clock c. An ancient sundial d. An ancient superstition |
| 2. Comprehension | <p>Which of the following best describes what was meant by "The past and present existing at the same time."?</p> <ul style="list-style-type: none"> a. The old light of Adromeda illuminates our earth today. b. Remains of the past are still evident today. c. Grandparents and grandchildren exist at the same time. d. Old rocks determine the age of the present. |
| 3. Application | <p>If there were no mechanical clocks around, how might you best determine the approximate hour?</p> <ul style="list-style-type: none"> a. By examining a tree trunk b. By geological explorations c. By looking at Stonehenge d. By looking at the sun |
| 4. Analysis | <p>Which of the following is <u>not</u> an assumption made by the film?</p> <ul style="list-style-type: none"> a. Time is constant b. Time is variable c. Time is defined d. Time is important |

Procedure and results. One form of the Multiple-Choice Test was assigned to each of three university students after the students had watched Part II of Reflections on Time. As before, it was expected that some subjects would correctly answer high level questions as well as the subsumed lower level questions, while others were expected to answer correctly only at lower cognitive levels.

The number of students correctly answering items at each of the cognitive levels did not support the expectations. As shown in Table 6, some subjects correctly answered a similar number of items across the Taxonomy, but not necessarily all items at lower subsumed levels. Moreover, one subject correctly answered all questions with the exception of one item at the Analysis level.

TABLE 6

Number of Correct Multiple-Choice Answers
at Each of the Six Levels of the Taxonomy
with a Maximum Score of 3 at Each Level

| Subject | Level | | | | | |
|---------|-------|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | 3 | 2 | 2 | 2 | 3 | 2 |
| 2 | 2 | 1 | 1 | 2 | 2 | 1 |
| 3 | 3 | 3 | 3 | 2 | 3 | 3 |

Discussion. The more objective scoring Multiple-Choice Test did not provide a better measuring device than the Short-Answer Test. Because an optimum level of response could not be identified it was still difficult to label a subject as functioning at a particular cognitive level based on the test data. As with the previous Short-Answer Test, the data threw doubt on the hierarchical nature of the Taxonomy, since those answering at higher levels did not correctly answer all items

at lower subsumed levels. Subjects also seemed to be approaching some "ceiling effect" where they answered most high level questions. This suggested that the questions were too easy for university students or that university students tend to attempt more difficult items, sometimes at the expense of more basic information.

At any rate, the Multiple-Choice Test based on the Taxonomy was not a satisfactory instrument for measuring cognitive skill. Although the number of subjects was small, the fact that parallel data for all subjects was obtained when using short-answer completion as well as multiple-choice items suggests that the small N was not responsible for the inconsistencies of the data. Also, because extreme care was taken to assure that items specifically represented each cognitive level of objectives, it seemed that something other than invalid items was contributing to the inconsistencies of the data.

Oral test items. In keeping with the purpose of investigating alternate forms of questioning, oral test items were also analyzed. The items were basically the same as those found in the Short-Answer Test. The only difference was the verbal rather than written presentation of questions and answers. Once again, in order to determine if a longer test was superior to a shorter one, a long form (having two items for each level) and a short form (having one item for each level) were used. In order to test as many questions possible there were two alternate series of questions for both forms. The items were arranged hierarchically according to the Taxonomy. These Oral Tests are found in Appendix A.

Procedure and results. After watching Part III of Reflections on Time, three university students experienced the short form of the test and six students experienced the long form. The investigator read each question to the subjects and allowed at least 10 seconds for them to begin responding. If no response was begun in that time, the next question was asked.

The number of items correctly completed on the short and long forms of the Oral Test are found in Tables 7 and 8, respectively.

TABLE 7

Number of Correct Answers at Each of the Six Levels of the Taxonomy with a Maximum Score of 1 at Each Level - Short Oral Test

| Subject | Level | | | | | |
|---------|-------|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 2 | 0 | 0 | 0 | 0 | 1 | 0 |
| 3 | 1 | 1 | 1 | 1 | 1 | 1 |

TABLE 8

Number of Correct Answers at Each of the Six Levels of the Taxonomy with a Maximum Score of 2 at Each Level - Long Oral Test

| Subject | Level | | | | | |
|---------|-------|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | 1 | 1 | 1 | 1 | 2 | 2 |
| 2 | 1 | 1 | 1 | 2 | 1 | 0 |
| 3 | 1 | 0 | 1 | 0 | 0 | 2 |
| 4 | 2 | 2 | 2 | 2 | 2 | 1 |
| 5 | 1 | 2 | 2 | 2 | 2 | 2 |
| 6 | 1 | 0 | 0 | 2 | 2 | 0 |

The number of subjects correctly answering items at each cognitive level did not support the expectation that those answering at high levels would also answer correctly at lower levels. Some answered a similar number of items across the Taxonomy while others answered more at the higher levels and still others seemed to approach a "ceiling effect" whereby they answered correctly almost all of the items.

Discussion. The Oral Test did not provide a good measure of cognitive skill. The subjectivity of judging the answers right or wrong could have contributed to the inadequacy of the instruments, but more than that, because of results with the various forms of items, suspicion was growing that the hierarchy of objectives was not valid. Additionally, the "ceiling effect" once again pointed to the possibility that university students may not have been the best sample to use since a variety of cognitive levels was desired, not just high cognitive levels at the expense of lower levels.

Discussion of results of preceding items forms. Taken together, the small samples used to explore each item form, i.e., short-answer, multiple-choice, and oral items, gave substantial evidence against the validity of the hierarchy of objectives within the Taxonomy. Some research which came to hand after these trials lends support to the conclusions concerning the hierarchy. Madaus, Wood, and Nuttall, (1973) investigated structural problems of the Taxonomy by assuming that, if the Taxonomy was valid, adjacent levels should have direct relationships but no direct relationships should be found between non-adjacent levels. Their results showed some indirect links between non-adjacent levels, implying discrepancies in the hierarchy.

Bearing in mind that the data from the present work could perhaps have been accounted for by these structural problems of the hierarchy, other complexities such as characteristics of the subjects may have also contributed to the inconsistencies. The subjects were university students who have, in general, been trained to answer any question put to them. With both the subjectively scored oral and short-answer items and the objectively scored multiple-choice items it was possible for such students to attempt all questions and to guess correct answers for a proportion of them.

Younger subjects, high school students, participated successfully in earlier work and are not overtrained to answer all questions presented to them. They are not so hesitant to admit they do not know the answer. Additionally, it has been shown that even very young children are capable of answering questions

representing high cognitive levels, as well as low levels (See Hunkins, 1972, p. 50; Sanders, 1966, p. 10). Since there was little doubt that younger subjects could exhibit the desired answering behaviours, such subjects were used in subsequent work.

Besides the complexities of the classification scheme and of the characteristics of subjects, the forms of items were considered. Those used thus far required answers of specific cognitive levels. All items demanded a specific level of answer by way of their construction. If a subject correctly answered any item he was said to be operating at a specific cognitive level. This was perhaps an unnecessarily artificial method of assigning a cognitive level to subjects since the level of answer was predetermined by the question presented. Asking a more open-ended question to which subjects were free to answer at any cognitive level may have been a more valid representation of the cognitive skill of an individual. Open-ended questions would not make a subject feel compelled to answer by guessing if he did not know the answer.

Open-ended questions. In order to rectify some of the previous problems the use of open-ended questions was explored using high school aged subjects. Open-ended questions such as "Tell me about it.", give no clue to subjects as to the level of answer required. Theoretically then, an answer representing any level could be given, and a subject is free to answer at the level of his own preference.

However, the consequence of such a strategy is to place the emphasis, by necessity, not on the classification of questions but on the classification of answers. The initial task then was to determine if open-ended questions could elicit answers which could then be classified according to the Taxonomy.

Procedure. Four subjects, aged 12 to 15 years, participated individually. Each viewed Part II of Reflections on Time and responded to open-ended questions such as "Tell me about the film." An attempt was made to elicit answers which represented many levels of the Taxonomy. By presenting many open-ended questions and probing those answers given, the subjects were given

opportunity to provide numerous answers, thus increasing the probability that they would at least touch all of the levels.

The first two subjects were also asked some of the previously used short-answer questions to determine if they were, in fact, capable of answering questions at all levels of the Taxonomy. Examples of the open-ended and short-answer questions asked of one subject are presented in Table 9.

TABLE 9
Examples of Open-ended and Short-Answer
Questions Asked of One Subject

| Open-ended | Short-answer | |
|---|---------------|---|
| Question | Level | Question |
| <p>Would you tell us about the film.</p> <p>Were you able to draw any conclusions about it?</p> <p>If it was trying to make people think, how did it do it?</p> <p>Why was it putting together so many different bits of information?</p> | Knowledge | <p>What evidence of time is supplied by nature?</p> <p>What was the most accurate clock mentioned in the film?</p> <p>What are some of the rhythms and patterns of time that primitive man recognized?</p> |
| | Comprehension | What did the film mean by the statement "A tenth of a second is a precious commodity." |
| | Application | What do you think the writer of the film would say if somebody asked him "What is time?" |
| | Analysis | <p>Can you identify any scene or bit of script that illustrates the unknown qualities of time.</p> <p>What were the main ideas in that film clip?</p> <p>Could you pick out some scenes in the film that illustrates man's awareness of time?</p> |
| | Synthesis | Could you define time in your own words? |

Tape recordings of proceedings were subsequently transcribed and the transcriptions used in the categorizing exercise.

Classifying answers to open-ended questions. From the transcripts of subjects' responses, lists of answers were compiled and classified by four independent judges. The answers were classified according to the Taxonomy as defined in Appendix C. The judges were required to carefully read each answer as supplied and to classify it according to the category it most closely represented.

The transcript from the open-ended questioning of one subject is found in Appendix B and an example of the classification of responses given by the judges is found in Appendix E. The classification as found in Appendix E is only an example of those given by the judges because many discrepancies between judges arose, making it virtually impossible to give a specific classification to each answer. The judges, working independently, interpreted the answers in a variety of ways, often classifying a single answer into two or more categories. There appeared to be too much room for personal interpretations of the answers and, therefore, of the classification of those answers. More objective coding procedures were needed.

In an attempt to get a greater degree of agreement between judges and in order to avoid confusion in how to classify an answer because of subtle semantic characteristics, some responses were standardized to the format "It was..." as found in Appendix E. These answers in standardized format were once again classified by the four independent judges. This time, the Taxonomy as defined in Appendix C was employed in an attempt to supply a more objective definition of the six categories.

The amount of agreement between the judges was still far from satisfactory. When the codings were compared to a criterion coding supplied by the investigator, there was total agreement for 4 out of 6 answers classified as Knowledge, 1 out of 6 classified as Comprehension, 1 out of 5 classified as Application, 0 out of 6 classified as Analysis, 2 out of 6 classified as Synthesis, and 2 out of 6 classified as Evaluation level answers. That is,

the percent of total agreement varied from 0% to 67%, indicating that a satisfactory amount of agreement between judges had not been reached.

"Results" and Discussion. In light of the low amount of agreement between judges, the definitions of the six categories of the Taxonomy did not appear to be clearly delineated enough for the purposes of the present study. This tended to strengthen the suspicion that had evolved in earlier work that the Taxonomy was inadequate for the intended purposes. Although the Taxonomy may be a useful tool for guiding teachers in constructing varieties of classroom objectives, it does not provide definitions precise enough to categorize the vast array of answers which may be given to open-ended questions.

Realizing that the data drawn from the open-ended questioning may be suspect because of low reliability, but in order to draw some meaning concerning the adequacy of using open-ended questions and high school age subjects, the classifications given by one judge only were analyzed. That is, although the four judges could not reach a consensus about the classification of many responses, by using an example of the classification of responses it was possible to draw some tentative conclusions about the use of open-ended questions.

As an example, the frequency of responses from one subject in each of the six categories of the Taxonomy (as classified by one judge) for both open-ended and short-answer questions are found in Table 10.

TABLE 10

Answers to Open-ended and Short-Answer Questions from One Subject, Classified According to the Taxonomy by One Judge.

| Level of Answer | Open-ended Question | Short-Answer Question |
|-----------------|---------------------|-----------------------|
| Knowledge | 4 | 2 |
| Comprehension | 2 | 1 |
| Application | 0 | 1 |
| Analysis | 0 | 3 |
| Synthesis | 3 | 1 |
| Evaluation | 2 | 2 |

According to this classification, the subject did answer correctly answers in each of the six categories when short-answer questions were asked and also gave answers corresponding to many of the categories when open-ended questions were asked. Based on the data it appeared that the subjects were, in fact, able to give answers which represented all six levels of the Taxonomy, and the open-ended questions demonstrated that the answers corresponding to many levels of objectives could be elicited by such questions. Data from other subjects implied that all categories could be elicited recommending further use of such questions.

However, although the subjects appeared to be capable of answering questions which represented objectives of all cognitive levels, the types of answers obtained could not be controlled by the examiner because of the nature of open-ended questions. Therefore, problems would have occurred with open-ended questions if a specific category of answer was required. That is, if a specific type of answer had been desired, the examiner's only recourse was to ask more open-ended questions in hopes that it would be manifested.

A Modified Taxonomy

In the preceding work, the Taxonomy had proved less than satisfactory. The major problems were (i) discrepancies between judges when coding completion items, and (ii) discrepancies in the expected hierarchical nature of responses, as shown most clearly in data from the multiple-choice test wherein subjects sometimes answered at higher cognitive levels but not at lower ones.

Difficulties in using the Taxonomy as a measuring device have arisen for others (See for example, Hunecke, 1970) and as well, the hierarchical nature of the Taxonomy has been queried by Madaus, Woods, and Nuttall (1973). Sanders (1966) also drew attention to the subjectiveness of the Taxonomy: ". . . teachers working with the Taxonomy of questions often disagree on the classification of a question. . .and all parties to the dispute can make good causes for their position (p. 7)." In the face of this situation it

became apparent that an alternative classification scheme would be necessary.

Developing an Alternative Scheme

The first step was to determine what categories of the Taxonomy were responsible for most discrepancies between judges and which they found easy to use. The judges' classifications of responses to the open-ended questions, both standardized and unstandardized forms, were analyzed for this purpose. Table 11 presents the total percent of agreement with a criterion coding for those answers.

TABLE 11

Percent of Total Agreement of Classifications
Between Four Independent Coders

| | Level of Response | | | | | |
|----------------------|-------------------|-------|---------|-------|------|-------|
| | Know. | Comp. | Applic. | Anal. | Syn. | Eval. |
| % of Total Agreement | 67 | 17 | 20 | 0 | 33 | 33 |

The data show greatest agreement with the Knowledge, Synthesis, and Evaluation categories and least agreement with the Analysis category. The judges stated that they found the Knowledge and Evaluation categories easiest to use (that is, they felt most confident about putting responses into those two categories) and also that the Synthesis classification was sometimes contained in other categories.

Based on the data from Table 11 and the statements from the judges, the answers were regrouped. An attempt was made to regroup them according to consistent semantic cues somewhat

within the general framework of the Taxonomy. The semantic cues hopefully provided for more objective coding and the Taxonomy provided some categories with reasonable agreement and bases on which to develop others.

Semantic cues were identified for the Knowledge and Evaluation categories first since these two categories appeared to represent two extremes on a continuum and were most easy to differentiate. The remaining answers were then analyzed as to consistent semantic cues with four additional categories developing. Bloom's Comprehension category was divided into two parts, one included statements inferred from the content of the film and the other included, in addition to the inferred statements, reasons for those inferences. The two new categories were called Comprehension and Understanding, respectively. The Application category was redefined as statements which applied what was presented in the film and/or mentioned the "effects" of the material. The Synthesis category seemed to be engulfed in other categories with no need for a separate classification. However, a number of speculative "if" statements remained which suggested the final category, Conjecture.

As a result, an alternative classification scheme, referred to as the Coding Schedule was developed and is summarized below:

1. Factualization - Memorization or recall.
2. Comprehension - Inference.
3. Understanding - Inference plus a reason.
4. Conjecture - Speculation.
5. Application/Usage - Statements of effect.
6. Evaluation - Judgmental statements.

The categories of the Coding Schedule were assumed to be independent and representative of most types of classroom answers and were not assumed to be hierarchical. Specific definitions as well as semantic cues for each category were formulated as presented in Appendix D. An example of the definition of one category and of the accompanying semantic cues follows:

Factualization - A Factualization response, as defined here, involves the memorization or recall of facts and processes. No inferences are made. This type of response is often subsumed in other categories. Cue words include: That, as in "I saw that the car was. . ." and "It showed that. . ."

Testing the Coding Schedule

In order to justify the use of the Coding Schedule, as opposed to the Taxonomy as an instrument for measuring the cognitive skill represented by subjects' answers, it was necessary to demonstrate that the Coding Schedule was a successful classification scheme in its own right and that it was a more reliable instrument than the Taxonomy. A comparison of results obtained by using both classification schemes was needed to determine which provided the most consistent results.

Procedure. Twenty volunteer teachers college students were used as judges in the comparison. They first watched Part II of Reflections on Time and were then randomly assigned either (i) the Taxonomy or (ii) the Coding Schedule, to use as the instrument for classifying a list of "answers". The Taxonomy was summarized from the Taxonomy of Educational Objectives (See Bloom, 1956, pp. 201-207) and the instrument is found in Appendix C. The Coding Schedule used is found in Appendix D.

Each judge was required to classify the same 15 "answers", such as: "It said that you can't have time without motion." (See Appendix E for a complete listing.) Ten judges used the Taxonomy and ten used the Coding Schedule, resulting in 150 "answers" classified according to each. The judges' classifications were subsequently compared to a criterion coding to determine the suitability of the schemes.

Results and discussion. Each judge, regardless of which scheme was used, was able to allocate every "answer" to a category. The classifications of the two groups (those using the Taxonomy and those using the Coding Schedule) were compared as to their percentage of agreement with a criterion coding supplied by the investigator. For the group using the Taxonomy there was a 29% agreement with

the criterion coding, and for the group using the Coding Schedule there was a 77% agreement with the criterion. (See Willemson, 1974, for a discussion about satisfactory amounts of agreement).

Given that the Coding Schedule was workable in that naive judges could readily use it for classifying typical "answers" and could do so with a reasonable amount of agreement, the grounds for preferring it to the Taxonomy seemed strong.

Using a Model

The next task was to pursue the implications of the newly derived classification scheme for the design and development of stimulus materials appropriate for the purposes of the study. Initially two problems presented themselves: (i) to structure the presentation of the visual stimuli and (ii) to build in the vicarious learning component.

The need to investigate vicarious learning necessitated a modelling situation wherein a participating subject would supply the necessary vicarious stimuli for an observer. The problem was how to elicit predetermined answers from a model in response to open-ended questions more efficiently than in previous work.

Open-ended questions were preferred over other item forms because they ultimately allowed for the free manifestation of a variety of answers. However, because of the low operant level of some categories of answers when employing such questions, the direct application of P-reinforcers for a specific category of answer was difficult. Due to the impracticability of eliciting specific categories of responses through the use of open-ended questions, it appeared impossible to reliably investigate the direct reinforcement of specific categories of answers and to investigate consistently modelled stimuli. However, by employing a model who followed a script, giving the necessary variety of answers to open-ended questions, and who appeared to be freely emitting responses, the effects of vicarious P-reinforcers could be examined. However this situation did not permit the examination of effects of direct P-reinforcers on the model.

Thus a trained model allowed (i) predetermined "answers" and (ii) accurate and consistent praise for specific types of "answers".

However, perusing the possibility of using a real life model to supply the stimulus material for observers, it became apparent that a videotaped model would provide more consistency from session to session. That is, although a real life model could provide consistent content in answers to open-ended questions, slight variations in appearance, voice inflection and other characteristics were not in control of the investigator. On the other hand, a videotaped model giving predetermined "answers" provided that added measure of consistency and control desirable in an experimental study. The videotaped model's performance could be replicated exactly for each subject.

Characteristics of the Model

Since earlier work pointed out difficulties in using university students as subjects and indicated the acceptability of younger subjects, the use of fifth-form (aged 15 to 16 years) subjects and model was proposed for ensuing work. Inasmuch as reinforcers are similarly effective for most members of a particular group (Bandura and Walters, 1963, pp. 10, 85) and students most often watch classmates responding to questions, it was desirable to have a model and subjects of similar ages.

Besides the age variable in determining an appropriate model, the sex of both the model and the subjects was a confounding variable in many vicarious learning studies (See for example Bandura and McDonald, 1963; Bandura, Ross, and Ross, 1969; Kanareff and Lanzetta, 1960a; Krumboltz and Schroeder, 1965; Krumboltz, Varenhorst, and Thoresen, 1967; McDavid, 1962). These studies show that males and females react differently to models and their behaviour is somewhat dependent upon the sex of the model. To control for the variance which could occur when using subjects and models of both sexes, it was proposed to use both a female model and female subjects in the experimental situation.

Choosing a P-reinforcer

Within a model-observer framework, it was then necessary to define specifically the stimuli used as the potential vicarious reinforcers. That is, if a videotaped model was giving a variety of "answers" to open-ended questions asked by someone taking the "teacher" role, that "teacher" was required to supply some stimulus that had the potential of functioning as a vicarious reinforcer. The stimuli that could be used as P-reinforcers needed to be defined objectively in order that valid representations could be supplied.

In the normal classroom situation a teacher often praises students for answers which meet her criterion for success, assuming that praise will strengthen the desired behaviour. Zahorik (1969) identified some responses to answers commonly made by teachers as simple praise, e.g. "good", "yes", and "that's good thinking" (p. 36). Such statements of praise can take on a reinforcing role - both directly and vicariously - as shown by research (See for example Allen, Henke, Baer, and Reynolds, 1970; Goetz and Baer, 1973; Hart, Reynolds, Baer, Brawley, and Harris, 1971; Broden, Bruce, Mitchell, Carter, and Hall, 1970; Ditrichs, Simon, and Greene, 1967; Kanfer and Marston, 1963).

Although praise has been used as an effective reinforcer, it should be noted that individuals have personal sets of phenomena and preferred activities which can be used as reinforcers (Premack, 1965). Accordingly, the ideal situation would be to find P-reinforcers for each individual to ensure efficient learning. However, it seems to be generally accepted that students prefer to be praised rather than not to be praised or reprimanded. Thus, although praise may not be the most potent reinforcer for each individual, it is one that is often generally effective for many students.

Since the present study was not dealing with specific individuals but was concerned with commonly used P-reinforcers, it was considered reasonable to employ statements of praise in an attempt to strengthen specified answering behaviours regardless of some writers' claims (See Kennedy and Willcutt, 1964) that individual reinforcement histories are a major contributor to confusion in the literature

about the effectiveness of reinforcers.

Depriving subjects of praise

In a regular classroom, a teacher usually acknowledges an answer to a question, whether or not it is the desired type of answer, and students generally expect some type of acknowledgement for their answers, whether they are right or wrong. The teacher praises the type of answer that was desired and acknowledges other answers with a supposedly neutral or negative statement. A neutral phrase tells the student that the answer was recognized but contains no value judgment of merit whereas a negative phrase tells the subject his answer was wrong.

A phrase of acknowledgement relates to the concept of deprivation in that a phrase of acknowledgement deprives a subject of praise. When the praise does occur it is then more effective as a reinforcer because the subject has been previously deprived. A neutral phrase, or even a negative phrase for that matter, must be sufficiently different from the P-reinforcer so that the subject knows when his answer is being praised and so that the subject is deprived of praise except when desired.

The advantages of depriving subjects of reinforcers are discussed by Travers, Reid, and Van Wagenen (1964). Discussing the implications of research on reinforcement, they point out that an experimental effect is more likely to occur if the subject has been deprived of the reinforcer for a period prior to the learning than if he has been satiated.

Also advocating the deprivation of reinforcers, Zahcrik (1969) found in his survey that infrequently used phrases by the teacher were better facilitators of behaviour than frequently used ones. He also concluded that more elaborate phrases had more reinforcing value than less elaborate phrases.

The present study accommodated to the above factors by:

- (i) using statements of praise as potential reinforcers;
- (ii) investigating the relative effectiveness of long (more elaborate) and short (less elaborate) statements of praise; and
- (iii) using a neutral statement to acknowledge all answers except those which were to be praised.

Summary of Developmental Work

The preceding discussion of the stages through which the study developed pointed out problems which were encountered initially, the solutions to those problems, and the new directions given to the study as a whole. In brief these comprised:

- (i) an attempt to operationalize the concept of cognitive skill by means of the Taxonomy of Educational Objectives which proved inefficient due to discrepancies in the coding of responses and doubts as to the truth of the hierarchy;
- (ii) trials to determine item characteristics in which closed questions, such as multiple-choice and short-answer items, were initially used and subsequently rejected (because they tended to predetermine the answers given) in favour of open-ended questions which avoided forcing answers;
- (iii) trials to determine an appropriate subject population that resulted in the rejection of university students as being too sophisticated and the selection of young (high school age) as being satisfactory because they were capable of giving answers which illustrated a variety of cognitive skills but did not attempt to answer every item;
- (iv) the development from the Taxonomy to a modified Coding Schedule which was found to be a successful and more reliable classificatory instrument;
- (v) the adaptation of the reinforcement framework for investigating vicarious reinforcement by using models and observers which proved inefficient because the investigator had to await the free manifestation of the desired category of answer from the model before praise could be given;
- (vi) the proposed use of a model on videotape who gave predetermined answers in all six of the categories of the coding schedule and being praised for one specific type of answer.

CHAPTER THREE

Procedure - Experimental Design

Introduction

The general reasearch strategy that emerged as a consequence of the previous trials implied a scenerio of events which would entail the use of individuals who, having been brought into the experimental situation, would:

- (i) view selected parts of the film Reflections on Time;
- (ii) watch a videotape in which a model was asked questions about the film and gave a variety of answers - no praise supplied;
- (iii) select one of the model's answers as being a preferred response category before treatment;
- (iv) watch a second videotape in which a model was asked questions about the film and gave a variety of answers, with praise being given for pre-determined types of answers;
- (v) be questioned in order to determine the effects of the vicarious praise.

Throughout the chapter reference will be made back to this scenario so that methodological procedures can be seen against the sequence of events that occurred. To facilitate this, these five phases have been given descriptive titles constructed to reflect the subjects' involvement in the situation:

- (1) Initial film viewing
- (2) First Videotape observation
- (3) Baseline question
- (4) Second videotape observation - with P-reinforcer
- (5) Second film viewing
- (6) Oral response interview

It is the purpose of this chapter to present the research hypotheses and to detail the specific procedures followed in:

- (i) developing stimulus materials; (ii) testing those materials in a pilot study; (iii) determining the optimum number of subjects; (iv) questioning subjects; and (v) coding subject's responses.

Research Hypothesis and Variables

In the description that follows, the research hypotheses have been put into two classes: (i) those dependent upon theory but also influenced by developmental work or by characteristics of the experimental situation, and (ii) those depending mostly upon an orthodox theoretical basis or related research.

Empirically Based Hypotheses

Early trials had shown that subjects gave answers to open-ended questions which could be classified into each of the six categories of the Coding Schedule. However, there was no reason to expect each category of response to appear equally in any given situation. More specifically, in response to open-ended questions, some types of answers may occur more often than others. For example, individuals may demonstrate their understanding of material more often than they issue value judgments concerning the material when freely responding to open-ended questions. Therefore, it was hypothesized that:

H₁: Answering is a function of cognitive predisposition: subjects' answers will tend to fall unevenly into categories rather than being dispersed equally.

Additionally, preceding developmental work based on written responses to lists of questions, implied that subjects may be influenced by a seriation effect wherein items at the beginning and end of lists are remembered better and completed more quickly (see for example Hovland, 1938; Jensen, 1962). It seemed that the order of presentation may influence what observers remember. Therefore, it was anticipated that the order of presentation of the model's responses would affect the answers given by the subjects and the hypothesis was formulated as:

H₂: Answering is a function of stimulusseriation: answer types given by the model at the beginning or at the end of the presentation will tend to be matched by the subjects more than those in the middle.

Theoretically and Research Based Hypotheses

Because research into classroom-related behaviour has shown that vicarious reinforcers have strengthening effects and because reinforcement theory postulates that stimuli may have strengthening powers, it was hypothesized that:

- H₃: Answering is a function of vicarious reinforcement: answer types for which the model was praised will tend to be employed by the subjects disproportionately more than will other answer types.

And alternatively, because reinforcement theory dictates the strengthening of responses preceding reinforcers rather than following them, it was hypothesized that:

- H₄: Answering is not a function of praise as an attention-getter: answer types, given by the model, which directly followed the praise will not be employed by the subjects disproportionately more than will other types.

Although a stimulus may serve an attention-getting function, and attention is an important factor in determining the amount of learning, the praise in the present work was directed explicitly at the answer which preceded it.

Background theory and research also implied that specific characteristics of the reinforcing stimuli influence the amount of vicarious learning which occurs (See for example, Bandura, 1969; Zahorik, 1969). Consequently, it was predicted that more intense praise would have a greater facilitating effect than less intense praise as would praise at the end of the videotaped sequence opposed to that in the middle.

- H₅: Answering is a function of stimulus characteristics: subjects will be influenced more by vicarious praise when the intensity of the praise is increased.

- H₆: Answering is a function of stimulus characteristics: subjects will be influenced more by vicarious praise when the praise occurs at the end of the observed sequence rather than in the middle.

The hypothesis concerning change in position from the middle to the end of the sequence is in keeping with the seriation effect which suggests that praise at the end of a sequence may be more effective because it is remembered better. And finally, since the controversy concerning the influence of awareness of contingencies of reinforcement¹ was left unsettled (See Kanfer, 1968), the present investigation aligned itself with Bandura's view that learning can occur without awareness, but awareness does facilitate the learning. Therefore, it was hypothesized that:

- H₇: Answering is a function of awareness of the contingencies of reinforcement: subjects who can be identified as aware of the contingencies of reinforcement will employ the praised category of response more often than those who are unaware.

Independent Variables

In order to investigate the above hypotheses, the following variables were manipulated:

- (i) the order of presentation of the model's responses - to detect a seriation effect if it existed;
- (ii) the category of answer that was praised - to detect vicarious reinforcement if it existed;
- (iii) the intensity of the praise given to the model - to detect effects of alteration of the modelling stimuli;
- (iv) the position of the responses that were praised - to detect effects of alterations of the modelling stimuli.

These variables were respectively labelled for convenience:

- (i) Order of Model's Baseline Responses;
- (ii) Praised Model Response;
- (iii) Intensity;
- (iv) Position.

¹ See Eriksen (1960), Galloway (1968), and Weinstein and Lawson (1963) for discussions concerning the difficulties in determining awareness.

Dependent Variables

The dependent variables used to test the hypotheses were:

- (i) Baseline Response Category - The category of answer (as classified by judges) given by a subject prior to treatment;
- (ii) Baseline Response Position - The position of the Baseline Response Category in the list from which it was chosen;
- (iii) Response Category - The category of answer (as classified by judges) given by a subject after treatment;
- (iv) Awareness - Awareness of contingencies of reinforcement as determined by oral questioning.

Developing Stimulus Materials

After defining the independent and dependent variables, attention was given to the development of stimulus materials to be used in an experimental situation. Prior to the experiment it was necessary to: (i) train judges who would subsequently aid in constructing a model's script; (ii) construct the script; and (iii) film a model on videotape. These procedures will now be discussed in that order.

Training Judges

Since subjects were required to observe a model giving "answers" representing all categories of the Coding Schedule in the first and second videotape observations (Phases 2 and 4) it was necessary to select appropriate "answers" for the model's script. Responses which were selected from answers given by subjects in earlier developmental work were classified by two trained judges (university students) as well as by the investigator in order to ensure a valid representation of the Coding Schedule.

A number of training sessions were used to familiarize the judges with the Coding Schedule. Initially, the classification scheme and responses representing each of the categories of the scheme were read by each judge. These items are found in Appendix D. The reasoning and disagreements concerning example responses

were discussed by the judges.

It was considered vital to have a self-explanatory scheme that could be used by untrained judges in the future. For this reason, a proliferation of verbal directions was avoided. In clarifying the definitions of specific categories, reliance was placed on feedback from the judges as to confusing categories or confusing examples so that an easy to use and objective scheme emerged. Only slight modifications of the original Coding Schedule were needed. Its final form appears in Appendix D.

The Script for the Model

The script for the model had to be designed so that, in response to the questions asked by the "teacher" in Phases 2 and 4, the model could appropriately produce "answers" representing the six different categories of the Coding Schedule. It was most important that truly representative "answers" for each category were supplied, with no disputes between judges, since observing subjects in the experimental situation were to see a model giving answers explicitly representing all categories, with one specific category being praised. To ensure that the classification given each "answer" was reliable, a procedure using inter-judge agreement was employed.

In order to have a sufficient supply of responses from which to construct the script for the first and second videotapes, three answers representing each category of the classification scheme were needed. To this end, each of the three judges classified a number of "answers" (presented in Appendix E) according to the final form of the scheme. Each "answer" began with the phrase "It was. . ." to standardize the format so that, in the operation of Phases 2 and 4, subjects would not be attracted to responses because of construction peculiarities.

Since there was a total agreement for only 16 of the first 18 responses coded by the judges, two additional ones were needed to supply three answers for each of the six categories of the classification scheme. Therefore, an additional 20 "answers" (See Appendix E) were classified. In this case, there was total

agreement between the three judges in 14 of the 20 classifications. In the six cases where there was disagreement, two judges agreed and a third did not. Two of the responses that had total agreement among the judges were chosen to complete the required 18 responses for the scripts. Complete scripts for all forms of the first and second videotapes are found in Appendix F.

The Studio Setup and Technical Considerations

For the production of the videotaped "teacher"-model interviews used in Phases 2 and 4, a studio was set up to simulate as far as possible the experimental setting that was to be used subsequently.

In the production of the videotapes, the model (unfamiliar to the subjects) wore a school uniform which represented standard clothing on the black and white videotape. The cameras were positioned so that after an establishing shot (showing a front view of both the model and the "teacher"), the remaining takes permitted a front view of the model giving "answers" with the back of the "teacher" in the foreground. Each question-answer episode was separated by a fade to grey. After the initial shot, the face of the "teacher" was not shown so that non-verbal facial cues would not be seen by the observing subjects since only verbal praise was to be used as a P-reinforcer.

Baseline Videotapes

To aid in measuring the category of answer a subject gave prior to exposure to vicarious praise, Baseline Videotapes were produced. These first videotapes were observed by subjects in Phase 2 of the experiment and showed the model giving six "answers" to open-ended questions asked by the "teacher". Each of the six "answers" represented a different category of the classification scheme, giving subjects exposure to all categories of answers.

Because there were grounds for believing that the order of the "answers" in the Baseline Videotape might affect those subsequently chosen by subjects, six different arrangements of the

same "answers" were employed. Each "answer" appeared at the beginning of one Baseline Videotape and at the end of another. The order of "answers" in the six Baseline Videotapes is found in Table 12.

TABLE 12

Order of Response Categories on Baseline Videotapes

| Position of Response Category | Baseline Videotape | | | | | |
|-------------------------------|--------------------|---|---|---|---|---|
| | A | B | C | D | E | F |
| First | 1 | 2 | 3 | 4 | 5 | 6 |
| Second | 5 | 6 | 1 | 3 | 6 | 4 |
| Third | 2 | 3 | 6 | 5 | 1 | 3 |
| Fourth | 3 | 4 | 2 | 6 | 2 | 5 |
| Fifth | 4 | 1 | 5 | 1 | 4 | 2 |
| Sixth | 6 | 5 | 4 | 2 | 3 | 1 |

Key to the Category of Response:

- 1 - Factualization answer
- 2 - Comprehension answer
- 3 - Understanding answer
- 4 - Conjecture answer
- 5 - Application/Usage answer
- 6 - Evaluation answer

For example, the first "answer" in Baseline Videotape A was classified as representing the Factualization category of the classification scheme. The sixth "answer" in that same videotape represented the Evaluation category.

One of the six Baseline Videotapes was assigned to each subject. In each, the model responded six times to an open-ended question about Part II of the film Reflections on Time, the same as was viewed by subjects in Phase 1 (initial film viewing). The open-ended question was: "(model), can you tell me about the film?" This question permitted the model to give one "answer"

representing each of the six categories of the classification scheme. Each "answer" given by the model was acknowledged by an unemotional and neutral-sounding "okay" from the "teacher." Since no praise was to be given in the Baseline Videotape, the "okay" deprived the model of praise while providing recognition of the "answer".

Treatment Videotapes

Treatment Videotapes were produced for the second videotape observation (Phase 4) to provide subjects with exposure to vicarious praise. In each Treatment Videotape the model gave twelve "answers" to open-ended questions about Part II of Reflections on Time. The same "answers" were used in each of the videotapes. Each of the six categories of the classification scheme were represented by two of the model's "answers" and "answers" representing only one particular category were praised in any of the Treatment Videotapes.

The six different Treatment Videotapes differed from each other in that; (i) each had a different ordering of the "answers" given by the model, and (ii) each showed a different category of response being praised. The six different arrangements of the model's answers allowed for the detection of a seriation effect and allowed the praised response to occur in comparable positions within each Treatment Videotape. The order of the answers in each of the six videotapes, along with an indication of the type of answer that was praised is found in Table 13. The complete model's script for each Treatment Videotape is found in Appendix F.

TABLE 13

Order of Response Categories on Treatment Videotapes

| Position of Response Category | Treatment Videotape | | | | | |
|-------------------------------|---------------------|---|---|---|---|---|
| | A | B | C | D | E | F |
| First | 5 | 6 | 1 | 2 | 3 | 4 |
| Second | 6 | 1 | 2 | 3 | 4 | 5 |
| * Third | 1 | 2 | 3 | 4 | 5 | 6 |
| Fourth | 2 | 3 | 4 | 5 | 6 | 1 |
| Fifth | 3 | 4 | 5 | 6 | 1 | 2 |
| Sixth | 4 | 5 | 6 | 1 | 2 | 3 |
| Seventh | 5 | 6 | 1 | 2 | 3 | 4 |
| Eighth | 6 | 1 | 2 | 3 | 4 | 5 |
| * Ninth | 1 | 2 | 3 | 4 | 5 | 6 |
| Tenth | 2 | 3 | 4 | 5 | 6 | 1 |
| Eleventh | 3 | 4 | 5 | 6 | 1 | 2 |
| Twelfth | 4 | 5 | 6 | 1 | 2 | 3 |

* Responses in these positions were praised

Key to Category of Response:

- 1 - Factualization answer
- 2 - Comprehension answer
- 3 - Understanding answer
- 4 - Conjecture answer
- 5 - Application/Usage answer
- 6 - Evaluation answer

From Table 13 it can be seen that there were two sets of responses corresponding to the classification scheme for each Treatment Videotape. That is, the first six responses (Set 1) represented all six of the categories, and the remaining six responses (Set 2) also corresponded to each of the categories. Also from the table, it can be seen that the third and ninth "answers" were praised. These positions were chosen so that the praise occurred near the middle of both sets of six responses which made up each Treatment Videotape. Thus when each subject was randomly assigned one of

the Treatment Videotapes, she was exposed to the same "answers" as all other subjects, only the order of "answers" and category of the praised response differed according to the particular treatment.

In each Treatment Videotape the model's "answers" were in response to the following open-ended question asked by the "teacher": "(model), can you tell me some more about the film?" All "answers", with the exception of those being praised, were acknowledged by an unemotional and neutral-sounding "okay" in order to recognize responses at the expense of praise. The phrase of acknowledgment, "okay", was sufficiently different from the praise, e.g. "That's a very good answer.", in order to assure that the observing subjects could distinguish between praised and not-praised "answers".

To allow the exploration of differential effects of various phrases of praise, one of two phrases were available and could be edited in for each Treatment Videotape. That is, either a long or a short phrase of praise could be used in any treatment Videotape. The short phrase ("That's a very good answer.") was similar for all response categories. The long phrase changed, depending upon the specific response category being praised. The long phrases used for each category are presented in Table 14.

TABLE 14
 Long Phrases of Praise Used
 for each Category of the
 Coding Schedule

| Category of Response being Praised | Long Phrase of Praise |
|------------------------------------|--|
| Factualization | That's a very good answer. You have remembered what the film said. |
| Comprehension | That's a very good answer. You have made an inference from what the film showed. |
| Understanding | That's a very good answer. You have made an inference from what the film showed and you have given a reason for the inference. |
| Conjecture | That's a very good answer. You have speculated about what you saw. |
| Application/Usage | That's a very good answer. You have applied what was presented. |
| Evaluation | That's a very good answer. You have made a value judgment concerning the film. |

Experimental Design

A description of the various phases of the experiment follows.

The Stimulus Film

Since both the Baseline and Treatment Videotapes showed the model giving answers based on Part II of Reflections on Time, common ground was first established by having each subject view the same film (Phase 1). The film was assumed to be equally unfamiliar

to all subjects.

Additionally, after they had seen both the Baseline and Treatment Videotapes (Phases 2 and 4), the subjects were shown a new part of the film (Part III of Reflections on Time) in order to tap a new domain of answers when testing for a treatment effect. Because they had to construct original answers to the investigator's questions (Phase 6) based on Part III of the film (Phase 5), the subjects could not simply reiterate answers given by the model. In earlier work a great variety of types of answers had been given in response to questions about Part III. Consequently, it was feasible to expect a similar variety of answers in response to questions about Part III asked of the subjects in the experimental situation. That is, the new stimulus material by no means limited the variety of answers which could be given. Both parts of the film allowed a similar diversity of answers to questions such as those asked of the model.

Baseline Response Category

After each subject had watched Part II of Reflections on Time (Phase 1) and a randomly assigned Baseline Videotape (Phase 2), a Baseline Response Category had to be deduced by the investigator - that is, the category of answer that the subject was predisposed to give prior to observing the selective praise of answers was determined. To do this, a written list of the model's "answers" in the order they were given in the Baseline Videotape, was furnished. From this, the subject was required to choose one of the "answers" as being the answer she would have given under similar circumstances to those of the model (Phase 3). The investigator determined the Baseline Response Category from this choice. The baseline for each subject was labelled as either Factualization, Comprehension, Understanding, Conjecture, Application/Usage, or Evaluation, depending upon which "answer" had been selected.

Detecting an Experimental Effect

In the experimental situation, if the effects of the subject's baseline tendencies were to be controlled, the praise given to the model on the Treatment Videotape had to be for a category of response different from the Baseline Response Category. To avoid experimental bias, after a subject's Baseline Response Category had been established, the investigator randomly assigned one of the five Treatment Videotapes that did not show the Baseline Response Category being praised.

It should be noted that in order to avoid wastage of subjects, Treatment Videotapes showing the long form of praise were used with the first set of subjects in the final experiment. Data from these subjects were examined to determine if it was profitable to explore a shorter and therefore weaker form of praise.

Subsequent to watching the randomly assigned Treatment Videotape and Part III of the film, each subject was questioned to determine if she was affected by the vicarious praise (Phase 6).

Since observers who detect some similarity between themselves and the model will presumably respond differently from those who do not (Bandura, 1969), both "personal" and "empathetic" questions were used. The Personal Question asked subjects how they would answer the question "Can you tell me about the film", while the Empathetic Question asked them how the model would answer.

Additionally, since awareness of the contingencies of reinforcement was one of the variables which might influence the amount of vicarious reinforcement (See Kanfer, 1968), subjects were also questioned to determine their awareness. The specific questions used to determine awareness were finalized in a subsequent pilot study.

Summary of Design

Using the components of the experimental design as described and operationalized by the stimulus materials, the following procedure emerged.

1. Subject watched Part II of Reflections on Time. (Phase 1)
2. Subjects watched a videotape of the model giving "answers" about Part II of the film with no praise for any "answers". (Baseline Videotape - Phase 2)
3. Subject selected one of the model's answers from which a Baseline Response Category was deduced. (Phase 3)
4. Subject watched a videotape showing the model being praised for one type "answer" about Part II of the film. (Treatment Videotape - Phase 4)
5. Subject watched Part III of the film. (Phase 5)
6. Subject responded to open-ended questions about Part III of the film in an attempt to detect a treatment effect. (Phase 6)

A Pilot Run

The next task was to formalize the emerging experimental procedures to be used in testing the effects of vicarious praise. This was accomplished by a prototype study in which the specific procedures and questions for subsequent experimental work would be finalized and difficulties with the operation of the videotape hardware would be identified and corrected without wastage of subjects. The general purpose of the prototype study, therefore, was to refine the experimental procedure.

Sample and Procedure

Six volunteer male subjects, aged 12 to 15 years, watched Part II of Reflections on Time and were assigned Baseline Videotapes and Treatment Videotapes which did not show their Baseline Response Category being praised. Three watched long phrases of praise and three watched short phrases of praise on the Treatment Videotapes. The specific Baseline and Treatment Videotapes (long or short praise) administered to each is indicated on Table 15.

TABLE 15
 Baseline and Treatment Videotapes
 Administered to Six Volunteers

| Subject | Baseline Videotape | Treatment Videotape | Length of Praise |
|---------|--------------------|---------------------|------------------|
| 1 | E | C | Short |
| 2 | E | C | Long |
| 3 | C | F | Short |
| 4 | C | F | Long |
| 5 | A | E | Short |
| 6 | A | E | Long |

Subsequent to watching the Treatment Videotape and Part III of the film, each subject was questioned to test those questions to be used subsequently in determining the treatment effects and awareness of contingencies of reinforcement.

Results and Discussion

No attempt was made to detect treatment effects since the sample was small and the general purpose of the pilot run was to test the emerging procedures for subsequent use in the experimental setting. From the pilot run it appeared that the proposed experimental procedure was satisfactory in that each step (Film - Baseline Videotape - Deduce Baseline Response - Category - Film - Treatment Videotape - Open-ended questioning) led easily from the preceding one and it was possible to quickly assess the Baseline Response Category and randomly assign a Treatment Videotape. Additionally, the operation of the Videotape equipment and the physical set up of the experimental room were found suitable for the purpose intended.

The questioning of these subjects was a proven ground for the oral questions to be asked of subjects in the ensuing experimental work. The two major questions used to determine an experimental effect were refined to:

1. What would you answer to the question, "Can you tell me about the film?"? - Personal Question
2. What do you think (model), the girl on the videotape, would answer to the question, "Can you tell me about the film?"? - Empathetic Question

Additionally, based on informal questioning about the Treatment Videotape, two questions were composed for the purposes of identifying subjects aware and not aware of the contingencies of reinforcement:

1. Did you notice any difference in the types of answers the (model) was giving?
2. Did you notice anything different in what I was saying to (model) on the videotape? Why do you think I said that?

Investigator's Role

Based on feedback obtained from the Pilot Run, the investigator's script for experimental sessions was formulated as found in Appendix F. The role of the investigator during experimental sessions involved the following:

1. The investigator gave a brief introduction to Part II of Reflections on Time and to the Baseline Videotape.
2. The investigator explained how the subject was to choose a Baseline Response Category from a written list of the answers given by the model on the Baseline Videotape.
3. The investigator assessed the Baseline Response Category and randomly assigned a Treatment Videotape, one which did not show the Baseline Response being praised.
4. The investigator gave a brief introduction to the Treatment Videotape.
5. The investigator then introduced Part III of the film by asking the subject to watch another piece of film about which she would subsequently be questioned.

6. After Part III of the film, a cassette recorder was started in order to have an accurate audio record of the questioning of subjects.

7. The investigator then asked the subject the Personal Question, the Empathetic Question, and questions to determine her awareness of the contingencies of reinforcement.

Optimum Number of Subjects

In any experimental treatment, parsimony is desirable. Not only should conditions permit the effective testing of hypotheses but they should also allow efficient use of subjects. Brewer (1972) states that it is desirable to determine an optimum number of subjects for a reasonable chance of detecting an experimental treatment effect if it is present. The optimum number of subjects, which would allow the detection of an experimental effect without wasting subjects, is dependent upon both the method of analysis (what statistical test is used) and the amount of chance error that would be tolerated in the decision of whether or not to reject a null hypothesis based on that method of analysis.

For this study, a number divisible by six was taken as a starting point for the number of subjects since this permitted an equal number of subjects in each of the six treatment groups. When an optimum number of subjects was determined for a reasonable chance of detection of an experimental effect, any additional volunteers could be used in variations of the original experimental treatments. For example, if 12 was found to be an efficient sample size, and after the treatment was carried out no effect was apparent using 12 subjects, then the use of a greater number of volunteer subjects would have added little information. Additional subjects could, more profitably, have been used in variations of the experiment based on the information gained from the first 12.

For the present study, the amount of error acceptable in rejecting a hypothesis of no experimental effect by chance was set at the customary 5% probability level. That is, 95% of the time we can be sure that a decision to reject a hypothesis of no experimental

effect will be correct, but there is a 5% chance that the wrong decision has been made, i.e. that an effect was wrongly detected. This chance of detecting an effect wrongly is referred to as alpha. By setting the alpha level, the experimenter sets the probability that he will reject a null hypothesis (a hypothesis of no treatment effect) because of a sampling error rather than because of a true treatment effect. The problem, then, is that of determining a suitable number of subjects (N) to allow for a reasonable chance of detecting an experimental effect when the experimenter is willing to accept a .05 probability level of chance error, Type I error or alpha error (See Senter, 1969, p. 185).

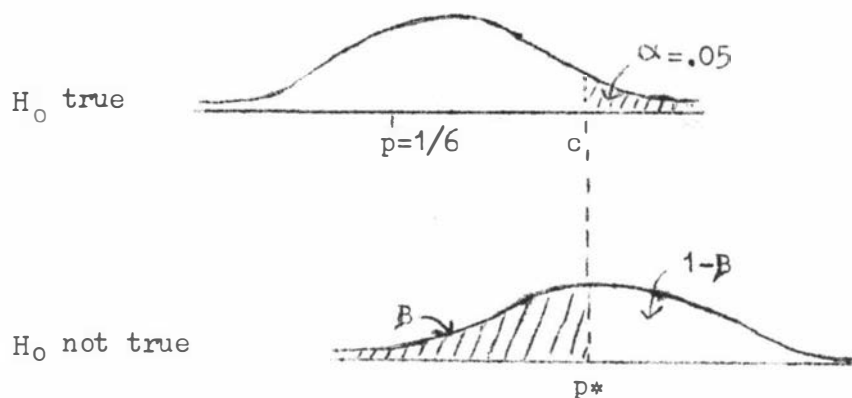
Type II or beta error occurs with the failure to reject a hypothesis of no experimental effect (H_0) when it is false. The probability of incorrectly failing to detect an experimental effect is represented by β . The power of a particular statistical test is the probability of correctly detecting an experimental effect (rejecting H_0 when it is false). For a given value of alpha, the power of a test ($1 - \beta$) increases as the sample size increases. By setting the power and the probability of alpha error to a reasonable level, the experimenter can thereby determine an optimum sample size (Brewer, 1972, p. 395).

The most representative example of a null hypothesis under consideration in the present investigation is the following: If a treatment effect did not exist (H_0 true), the probability is $1/6$ ($p=1/6$) that the subjects' answers will match those which were praised. If a treatment effect did exist (H_1 true), the probability was greater than $1/6$ ($p > 1/6$) that the subjects' answers matched those which were praised. In other words: (Null hypothesis) $H_0: p=1/6$ and (Alternate hypothesis) $H_1: p > 1/6$. Any value of $p > 1/6$ may be represented by p^* . If a treatment effect exists, and $p=1/6$ is not true, some other p^* value is true.

Figure 1 illustrates the conditions under which an existing treatment effect may be detected.

FIG. 1.

Conditions under which a Treatment Effect
may be Detected



- α : Probability of rejecting H_0 when it is true
 β : Probability of not rejecting H_0 when it is false
 $p=1/6$: Proposed value of p
 p^* : True value of p
 $1-\beta$: Probability of rejecting H_0 when it is false
 c : Value corresponding to the critical region for α

The probability of not rejecting H_0 when it is false (that is, $p > 1/6$) is represented by β and the area under the curve thus labelled. This is the probability of incorrectly failing to detect a treatment effect. The probability of correctly detecting a treatment effect is thus represented by $1-\beta$ (the power of the test), the remaining area under the curve. In general, the larger the difference between the actual value of p^* and the chance of $p=1/6$, the easier it will be to detect an experimental effect and, therefore, the greater will be the power of the test.

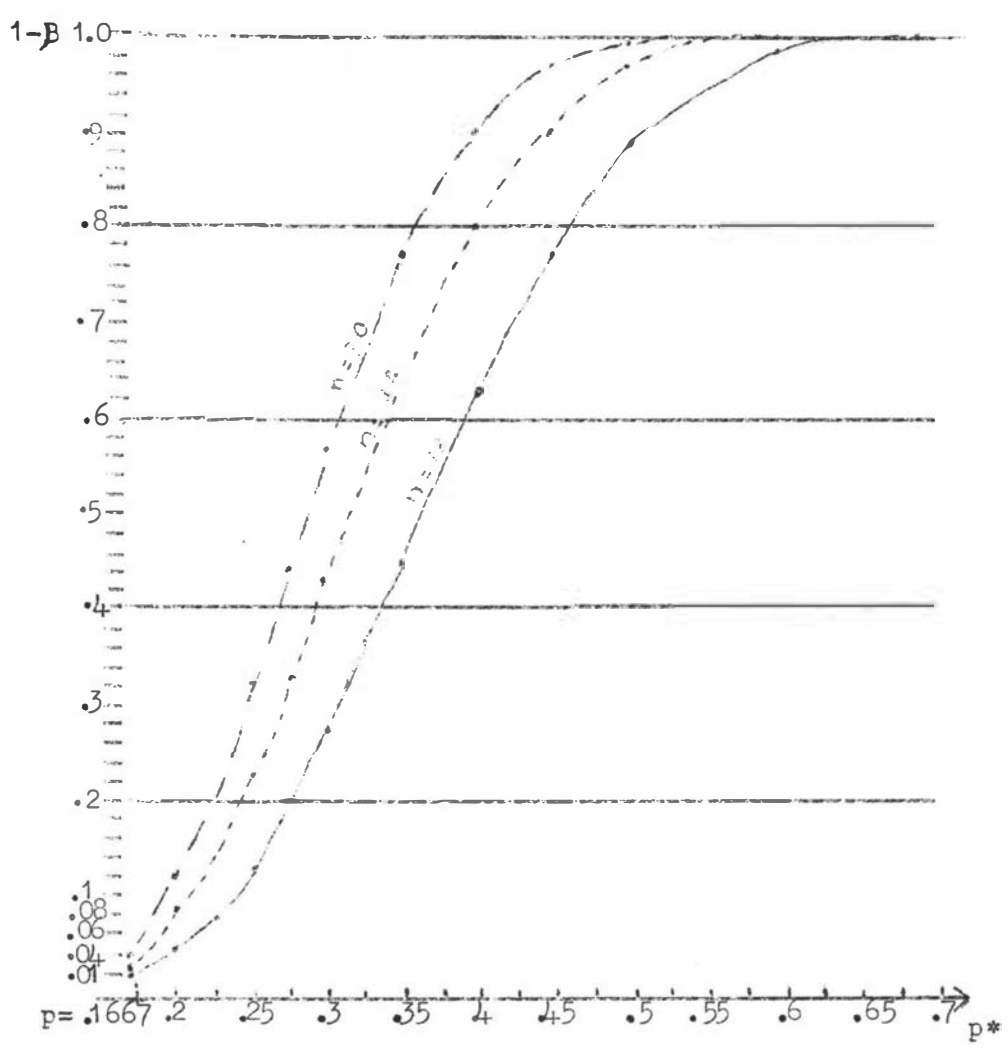
For a specified number of subject, a power curve can be drawn from which the probability of correctly rejecting H_0 for any value of p^* can be determined (See for example Glass and Stanley, 1970, pp. 283-288). The power ($1-\beta$) is dependent upon the value of p^* and the cut-off point for the critical region (represented by c), which depends on the value of α . The power is calculated using the binomial expansion

$$1-\beta = \sum_{x=c}^N \binom{N}{x} p^{*x} (1-p^*)^{N-x}, \text{ where } N = \text{total number of}$$

subjects and c =minimum number of subjects required to show an effect when $\alpha \leq .05$ and $p=1/6$. For each N the critical number of subjects needed to show a treatment effect in order to obtain significant results was determined by letting $p=1/6$ and calculating exact probabilities for each N (Appendix I presents the binomial expansion for each N with the critical cut-off point indicated in each case). For $N=18$, if six or more subjects showed a treatment effect, the results would have been significant below the customary .05 level. For $N=24$, if seven or more subjects showed a treatment effect, the results would have been significant below the .05 level. For $N=30$, the cut-off point fell right at eight, so if eight or more subjects showed a treatment effect, the results would have been significant at the .05 level.

The proposed values of N were multiples of six to allow for subjects to be assigned to each of the six treatment groups. To plot each curve, the value of p^* was varied from .1667 to .5000 and the resulting values of $1-\beta$ were plotted. Appendix I presents the computer programme that was written to calculate the binomial probabilities for the power values. A family of power curves for values of $N=18$, 24, and 30 were plotted to compare the power of the statistical tests with increasing N 's. The respective values of six, seven, and eight as alpha intercepts were used for $N=18$, 24 and 30, and by varying the values of p^* , $1-\beta$ was calculated. Values of p^* gave points along the power curve of $1-\beta$ for power curves of $N=18$, 24, and 30 as shown in Figure 2.

FIG. 2
Power Curves For N= 18,24,30



| p * | 1-β | | |
|-------|------|-------|-------|
| | n=18 | n=24 | n=30 |
| .1667 | .021 | .035 | .050 |
| .2000 | .051 | .089 | .129 |
| .2500 | .139 | .234 | .326 |
| .2750 | .203 | .330 | .447 |
| .3000 | .278 | .435 | .568 |
| .3500 | .451 | .642 | .775 |
| .4000 | .626 | .808 | .906 |
| .4500 | .774 | .914 | .969 |
| .5000 | .881 | .968 | .992 |
| .6000 | .980 | .998 | 1.000 |
| .7000 | .999 | 1.000 | |

As an example, for $N=18$, if the actual value of $p^* = .2750$ is compared to the chance value of $p = .1667$, the curve for $N=18$ shows the probability to be .20 that a false H_0 would be rejected. That is, the probability of correctly detecting an experimental effect would be .20. Since the value of p^* in this case does not differ much from the chance value of p , the test is not very powerful, allowing correct detection of the experimental effect only 20% of the time.

A more reasonable value of p^* to expect if there was, in fact, an experimental effect was $p^* \geq .40$; that subjects' answers would match those that were praised 40% or more of the time. If it is reasonable to expect the detection of an effect in which matching occurred 40% rather than 17% ($p = .1667$) of the time, we would want a high probability of detection, say at a level approaching .90. An effect that caused matching much less than 40% of the time would not be strong enough to warrant the same degree of interest especially since, by chance, matching would occur 17% of the time with no experimental effect.

If the actual value due to an experimental effect was $p^* = .40$, the probability of making the right decision in rejecting H_0 for $N=18$ would be .63, for $N=24$ the corresponding probability of making the right decision in rejecting H_0 would be .81, and for $N=30$ it was .91, a very high probability of detecting an experimental effect as high as $p^* = .40$. Thus, $N=30$ allowed for a reasonably powerful test when attempting to detect an experimental effect and, therefore, 30 subjects were considered an efficient sample size. This permitted 5 subjects in each of the six treatment groups which observed one specific category of response being praised. Additional volunteer subjects could then be used in alterations or initial treatments.

Sample

Based on the analysis of the power of the statistical test, 30 was deemed a satisfactory and efficient N . Fifth-form subjects (aged 15 to 16 years) were used because the film was aimed primarily at a high school audience and high school students had proved satisfactory in initial work. Volunteer female subjects were sought from high schools in a New Zealand city.

Sample Acquisition

Three high schools were visited in the initial contact, and appeals for volunteers were made during morning assemblies. Although the use of volunteer subjects and non-random sampling somewhat limits the generalizability of the results of the present study, it was the most feasible way of obtaining willing subjects. An attempt was made to contact each fifth-form girl in public school in a small New Zealand city in order to obtain a wide variety of fifth-form girls.

The volunteers who agreed to participate were subsequently contacted at home by telephone to confirm parental permission and to set a convenient time for the one-to-one treatment session¹, resulting in 65 appointments. Although measures were taken to assure that all volunteers would participate, i.e. reminder notes and telephone calls were made, only 53 actually experienced an experimental treatment, with the other 12 dropping out for reasons unknown.

Treatment Videotapes used with the Sample

The first 31 subjects (one was later eliminated because she did not respond to either the Personal or Empathetic Questions) experienced the experiment as it was originally designed. The long phrase of praise was used with each of these subjects and 5 subjects observed the praise of one particular category of response. The remaining 22 subjects experienced alterations of the original design - 11 were administered a more intense praise and 11 were administered praise at the end of the list of responses rather than in the middle.

Specification of Treatments and Nature of Responses

Testing the effect of vicarious praise in the experimental setting, subjects were individually exposed to stimulus materials and questions to determine treatment effects. The videotapes observed by each subject are now listed and uses made of their responses are presented.

Specification of Videotapes

The specific Baseline and Treatment Videotapes assigned to each subject is indicated in Table 16.

¹ The treatment session for each subject was in the neighbourhood of 35 minutes long and required a scheduled block of one hour.

TABLE 16

Baseline and Treatment Videotapes
Randomly Assigned to Subjects

| Subject | Baseline Videotape | Treatment Videotape |
|---------|--------------------|---------------------|
| 1 | D | D |
| 2 | B | E |
| 3 | D | F |
| 4 | F | F |
| 5 | D | C |
| 6 | E | C |
| 7 | E | A |
| 8 | A | E |
| 9 | C | C |
| 10 | A | C |
| 11 | A | C |
| 12 | F | D |
| 13 | E | A |
| 14 | A | D |
| 15 | C | B |
| 16 | F | B |
| 17 | D | A |
| 18 | B | B |
| 19 | F | E |
| 20 | B | A |
| 21 | C | D |
| 22 | C | A |
| 23 | E | B |
| 24 | B | E |
| 25 | B | D |
| 26 | F | F |
| 27 | E | E |
| 28 | A | F |
| 29 | D | F |
| 30 | C | B |
| 31 | C | D |
| 32 | A | A |
| 33 | C | B |
| 34 | C | E |
| 35 | E | A |
| 36 | F | A |
| 37 | F | E |
| 38 | B | A |
| 39 | D | E |
| 40 | D | A |
| 41 | A | A |
| 42 | B | E |
| 43 | C | E |
| 44 | D | A |
| 45 | C | E |
| 46 | E | A |
| 47 | E | E |
| 48 | B | A |
| 49 | A | A |
| 50 | F | E |
| 51 | B | A |
| 52 | F | E |
| 53 | A | A |

Long phrase
of Praise

Change in Position
of Praise

Change in Intensity
of Praise

The Baseline Videotapes, lettered A to F, indicate six different arrangements of the model's "answers" on the videotape whereas the Treatment Videotapes, lettered A to F, indicate a specific ordering as well as the specific category of answer that was praised. The praise of Factualization, Comprehension, Understanding, Conjecture, Application/Usage and Evaluation is represented respectively by Treatment Videotape A, B, C, D, E, and F. Each Baseline and Treatment Videotape was randomly assigned to five subjects.

Measures taken on subject were derived from:

- (i) The Baseline Response Category
- (ii) Responses to the Personal and Empathetic Questions
- (iii) Responses concerning awareness of contingencies of reinforcement.

These responses and their uses are now discussed in that order.

Baseline Response Category

The Baseline Response Category was assessed quickly and objectively from the subject's choice on a list of the corresponding answers which had been given by the model on the Baseline Videotape. These data were used in three ways: (i) to determine which Treatment Videotape could be randomly assigned to the subject; (ii) to determine if the ordering of the answers on the Baseline Videotape affected the choice made by the subjects; and (iii) to determine if subjects chose some categories more often than others when no praise was offered to the model.

Responses to the Personal and Empathetic Questions

Data from the oral questions which followed the experimental treatment were more complex than that resulting from the Baseline Response Category. After the cassette tapes were accurately transcribed, each response to the Personal and Empathetic Questions were coded as to the category of answer they represented. These coded responses were then used to determine if the vicarious praise affected the type of answers given. A representative transcript from Subject 4 is found in Appendix G.

Most subjects (all but one) responded to the Personal Question but many chose not to answer the Empathetic one. The answers to the Personal Question were coded as to which of the six types of

categories were present - most answers had more than one category present. In the analysis of the data concerning the Personal Question, both the first code (Initial Response Category) of a subject's answer and the code that was present most often (Representative Response Category) were used to determine if subjects matched the category of response that was praised more often than was expected by chance alone. Both the Initial and Representative Response Categories were determined for answers to the Empathetic Question, but the resultant codings were identical.

The code of the subjects' responses to the Personal and Empathetic Questions were said to either match or not match the category that was praised on the Treatment Videotape, depending upon whether the subject's response was coded the same as the category that was praised or not. The probability of the observed number of matches was determined to test the significance of results when praise was experienced vicariously.

Determining Awareness

Subjects were asked two questions to measure their awareness of the contingencies of reinforcement. These questions required subjects to (i) state the differences in the types of answers given by the model, and (ii) state the praise and the type of answer that was praised. If a subject could state the "answer" that was praised and was also able to explain why that type of answer was different from others, although all six categories may not have been accurately identified, the subject was considered to be aware of the contingencies of reinforcement. The data were used to determine if subjects' answers were a function of the awareness.

Coding Responses

Data from the Personal and Empathetic Questions were coded by two trained judges and the investigator. Such data were unobtainable from one subject because of no codeable responses to the questions. The investigator was able to deduce a Baseline Response Category for this uncodeable subject, but when verbal answers were required, as was the case for the Personal and Empathetic Questions as well as for those used to determine awareness of the contingencies, the uncodeable subject did not

respond. Therefore, this subject was eliminated from the analyses concerning verbalized responses.

Each of the remaining 52 subjects' answers to the Personal and Empathetic Questions were divided into codeable units since many included more than one sentence or phrase and, hence, more than one category of answer. The responses were broken down into codeable units prior to the coding in order to have the same number of codes from each judge allowing for more precision than if the judges themselves divided the responses into units.

Inter-Judge Consistency

From the codings for the Personal Question, there was total agreement between the three judges (referred to as A, B, and C) 61% of the time. From the codings for the Empathetic Question, there was total agreement 62% of the time. Judges A, B, and C agreed with one or both of the others 74%, 85%, and 94% of the time, respectively, when coding responses to the Personal Question; and 68%, 84% and 89% of the time, respectively, when coding responses to the Empathetic Question. When total agreement was lacking as to the code of a response, the code on which two of the judges agreed was used and in the 14 cases where all three judges disagreed, a consensus coding was used. Appendix H presents the codings which developed from this procedure.

Intra-Judge Consistency

In order to establish intra-judge reliability, one month after the first coding of the data, each of the three judges re-coded the 161 units of response. For each judge, the first and second codings were compared using a Pearson Product-Moment Correlation. The correlations were .78, .75, and .92 for Judges A, B, and C, respectively.

Analyzing Disagreements

It appeared that some coding categories were more responsible for inconsistencies between judges than others. In an attempt to investigate these discrepancies in order that problem categories might be identified for alternative examinations of the data, analyses were done to determine which categories of the Coding

Schedule caused the most discrepancies between judges.

For the Personal Question, 68% of the disagreements in coding involved the Factualization category (Coded 1). That is, 68% of the time, a disagreement occurred when one or two of the judges classified the answer as a 1 but the others did not. Also for the Personal Question, the percentage of disagreements involving the remaining categories varied from 2% to 64%; and for the Empathetic Question, the percentage of disagreements involving each of the categories varied from 0% to 86% with the Factualization and Comprehension categories causing most disagreements (See Table 17).

TABLE 17
Percentage of Disagreements in Coding
involving Six Categories over Two Questions

| Category | Personal Question | Empathetic Question |
|----------|-------------------|---------------------|
| 1 | 68% | 71% |
| 2 | 64% | 86% |
| 3 | 27% | 7% |
| 4 | 18% | 30% |
| 5 | 23% | 28% |
| 6 | 2% | 0% |

For both questions, Code 1 and Code 2 were responsible for most discrepancies. This suggests that these two categories needed more clarification or needed to be combined into the same category. Disagreements involving only Code 1 and Code 2 accounted for 32% of the disagreements for the Personal Questions. That is, one or more of the judges gave the answer a code of 1 while the others gave it a code of 2 for 32% of the disagreements. For the Empathetic Question, disagreements involving only Code 1 and Code 2 accounted for 26% of all disagreements. Code 1 and 2 answers

were hard for judges to differentiate. This fact, in conjunction with evidence from subjects' responses about the types of answers given by the model, provided the basis for subsequent alterations in the analyses of the data wherein collapsed categories were used.

Additional Code Assignment Rules

Because of the splitting of responses into codeable units, certain adjustments to the assignment of codes were needed. That is, since the definitions of Comprehension and Understanding responses were similar in that each required inferences and different in that Understanding responses required reasons for the inferences, care was taken to assure that responses were classified as Understanding (Code 3) when a reason was supplied for a specific inference and as Comprehension (Code 2) when no reason was supplied. Therefore, if two adjacent codeable units of an answer were coded 2 and 3, and final code of 3 was assigned to the response if the unit coded 3 supplied a reason for the unit coded 2. Thus, a Comprehension answer could stand by itself or be subsumed in an Understanding answer. An adjustment in coding because of a Comprehension unit being subsumed in an Understanding answer was a relatively minor case which occurred only once.

The construction of the Empathetic Question also made certain adjustments necessary. Many subjects, 11 of the 19 who responded, began their responses with a conjecture such as, "I think she would have said. . ." The code that was used in the analyses, therefore, was the code of the part of the answer which followed the conjectural (Code 4) introduction. For instance, a response to the Empathetic Question with codes of 4 and 5 was given the code of 5 for the final analyses.

See Appendix H for a complete listing of the codings made by the three judges, taking the above rules into consideration, and the Initial and Representative Response Codes that were derived from those codings.

Summary

An experimental situation wherein subjects watched a videotaped model giving many types of answer to open-ended questions was devised to explore the effect of vicarious praise on cognitive skill exemplified by specific answering behaviours. A videotaped model provided a supply of consistent and predetermined responses for the observers. Two videotaped sequences were filmed - one in which the model gave "answers" which corresponded to each category of the Coding Schedule without being praised for any (the Baseline Videotape) and one in which a particular type of "answer" was praised (the Treatment Videotape). The material on which both the model's "answers" and the subsequent questioning of the subjects was based came from a film, Reflections on Time.

The effects of the vicarious praise shown on the Treatment Videotapes, was determined by comparing a subject's category of answer after the Baseline Videotape to that after the Treatment Videotape. In order to explore the effects of the vicarious praise and to determine the effects of alterations of the modelling stimuli and of awareness of the contingencies of reinforcement, the following variables were manipulated:

Independent Variables

1. Order of Model's Baseline Responses - The order of presentation of the model's responses was varied.
2. Praised Model Response - The category of answer that was praised was varied.
3. Intensity - The intensity of the praise given to the model was varied.
4. Position - The positions of the responses that were praised were varied.

Those variables which changed as a result of manipulation of the Independent Variables are as follows:

Dependent Variables

1. Baseline Response Category - The category of answer (as classified by judges) given by a subject prior to treatment.
2. Baseline Response Position - The position of the Baseline Response Category in the list from which it was chosen.

3. Response Category - The Category of answer (as classified by judges) given by a subject after treatment.
4. Awareness - Awareness of contingencies of reinforcement as determined by oral questioning.

A pilot run was used to finalize the experimental procedure, to test the operation of the technical equipment, and to formalize the questions used to determine a treatment effect and those used to determine awareness of the contingencies of reinforcement.

In order to detect an experimental effect and to make best use of available volunteer subjects, an efficient sample size was calculated. This was based on the power of the statistical test to be used in the analysis of the data. An efficient sample size of 30 was calculated, so additional volunteers could be used in alterations of the original experimental design, that is, using variations of the position and intensity of the praise.

The subjects were volunteer fifth-form females, a total of 53. The original design was used with the first 31, one was eliminated because of no responses to questions which required verbal answers, and alterations of intensity and position of the praise were used with the remaining 22 subjects. The data were coded by three independent judges according to specific coding procedures.

CHAPTER FOUR

Results

Introduction

It is the general purpose of this chapter to present the statistical analyses used to test each major hypothesis and to discuss the development of minor hypotheses with their accompanying analyses. To this end, analyses used to test each major hypothesis are examined in turn, followed by a discussion of the rationale and statistical analyses for the minor hypotheses.

Hypothesis 1

Hypothesis 1: Some Categories of Answers will be Chosen More Often Than Others When No Praise is Given

As an initial step in the experimental procedure, subjects, having viewed a randomly assigned Baseline Videotape, selected one of the model's "answers" as being similar to what they would have answered under comparable circumstances. The model's "answers" were classified according to the six categories of the Coding Schedule: (i) Factualization; (ii) Comprehension; (iii) Understanding; (iv) Conjecture; (v) Application/Usage; and (vi) Evaluation. This classification predetermined the category of the responses given by the model, one of which was selected by subjects as a baseline response. The distribution of Baseline Response Categories for all subjects appears in Table 18.

TABLE 18
Distribution of Baseline Response Categories

| Baseline Response Category | 1 | 2 | 3 | 4 | 5 | 6 | Total |
|----------------------------|---|----|---|----|---|---|-------|
| Frequency | 9 | 19 | 3 | 11 | 9 | 2 | 53 |

Table 18 shows that the baselines were distributed somewhat unevenly over all six categories with Category 2 (Comprehension) carrying approximately 36% of the cases, Category 4 (Conjecture) carrying 21%, and Categories 3 and 6 (Understanding and Evaluation) something less than 6% of the cases each.

In order to investigate Hypothesis 1, the frequency of subjects selecting "answers" which were classified as belonging to each of the six categories was compared to the frequency which would be expected to occur by chance alone.

It should be noted that, in the present work, where possible, exact probability calculations were used to test for the presence

of a treatment effect. However, where analyses with more than one degree of freedom were appropriate, chi square techniques were used to make inferences about the frequency of responses observed. In this case, a chi square statistic with five degrees of freedom was calculated using an expected frequency of $1/6$ of the total responses for each of the six categories of answers. With $N=53$, the expected frequency for each of the six categories was $1/6$ of 53 or 8.833, since the six categories could be expected to be divided equally over the 53 subjects. The chi square calculated ($\chi^2_5 = 21.378$, $p < .05$) was significant at the .05 level. Therefore, the null hypothesis that all categories were equally likely to occur by chance was rejected. This implies that "answers" of some categories were chosen more often than others. Under normal conditions, one would not expect to find all of the categories equally represented; rather, answers of some categories would be used more often than others.

The observed frequencies of selected "answers" in each category were subsequently used as best estimates of the expected baseline frequencies in later analyses. Without the present information about the frequencies of answers chosen in each category, the best estimate for each category would be simply $1/6$ of the total. However, the data obtained herein show that each category of answer is not equally likely to occur and were, therefore, used to calculate more precise estimates of the true frequencies.

Hypothesis 2

Hypothesis 2: The Order of Presentation of the Model's Answers Will Affect Subjects' Answers

To determine if the order of presentation of the model's "answers" affected subjects' responses, data from both baseline and treatment measures were investigated. For a baseline measure, it was possible for a subject to select any one of the model's "answers" as being closest to what she would have answered under similar circumstances. Therefore, a subject could select an "answer" in either the first, second, third, fourth, fifth, or sixth position of the specific Baseline Videotape which had been watched, regardless of the category of that "answer".

For measures taken after the treatment, i.e. responses to both the Personal and Empathetic Questions, a subject was classified as giving a particular category of answer. That answer was then labelled according to the position of that particular category on the Treatment Videotape.¹

Hypothesis 2 was tested using baseline and treatment measures. First, the number of subjects that actually chose "answers" in each of the six positions of the Baseline Videotape was compared with the number expected if answers in each of the six positions were equally likely to be chosen. Next, the number of subjects labelled as corresponding to each position of "answer" from the Treatment Videotape was compared with the numbers expected if responses were unaffected by the position on that Videotape. The labelling of the subjects' answers was done for each of the following dependent variables: (i) the Initial Response Category to the Personal Question, and (ii) the Representative Response Category to the Personal Question. These three tests of Hypothesis 2 will now be presented.

¹ Each Treatment videotape showed two sets of "answers" corresponding to each of the six categories. Moreover, the order of those responses, according to which category was represented, was the same in both sets of "answers". Since a response category held identical positions in both sets comprising a Treatment Videotape, the position of each response category was labelled according to its place in any one of the two sets of "answers", i.e. first through sixth.

Analysis of Baseline Measure

The distribution of the frequencies of those classified as selecting answers in each of the six positions of the Baseline Videotape, depending upon which Baseline Videotape had been watched appears, in Table 19.

TABLE 19
Distribution of Position in Baseline
Videotape of Baseline Response Categories

| Position of Baseline Response Category | 1 | 2 | 3 | 4 | 5 | 6 | Total |
|---|----|---|---|----|----|----|-------|
| Frequency | 10 | 4 | 7 | 10 | 12 | 10 | 53 |

It can be seen from the data presented in Table 19 that the baselines were somewhat unevenly distributed over the six positions of the Baseline Videotape. "Answers" in position 5 were selected most often (25% of the cases), those in positions 1, 4, and 6 were selected in approximately 19% of the cases, and those in positions 2 and 3 were selected in approximately 8% and 13% of the cases, respectively.

These observed data were compared to frequencies that were expected to occur by chance alone using a chi square statistic. The expected frequency of subjects selecting "answers" in any one of the six positions of the Baseline Videotape was $1/6$ of the total responses, $N=53$, or 8.833. Although the choices appeared to be somewhat unevenly distributed over the six positions, the calculated chi square, with five degrees of freedom, was not significant at the .05 level ($\chi^2_5 = 4.46422$, $p > .05$). The null hypothesis that the order of the model's "answers" would not affect subjects' responses was not rejected. This implies that the order of presentation of "answers" on the Baseline Videotape did not significantly affect the Baseline Response Categories chosen by subject. No seriation effect, wherein answers at the beginning

or end of a list are chosen more often than others, was evident.

Initial Response Category

After viewing a Treatment Videotape that showed the praise of a specific type of "answer" given by the model, each subject was asked how she would respond to an open-ended question similar to that asked of the model. The subject's initial response to this so-called Personal Question was classified by position, depending upon the particular position of the Treatment Videotape of the like category of response. This analysis was designed to determine whether the positions of the model's "answers" influenced the type of answer given by subjects.

The distribution showing the frequency with which the category of the subjects' initial response corresponded to each of the six positions in the Treatment Videotape appears in Table 20.

TABLE 20

Distribution of Position in Treatment
Videotape of Initial Response Category
to the Personal Question

| Position of Response Category | 1 | 2 | 3 | 4 | 5 | 6 | Total |
|-------------------------------------|---|---|---|---|---|---|-------|
| Frequency | 5 | 2 | 8 | 7 | 4 | 4 | 30 |

It can be seen from the table that the frequencies of responses corresponding to categories in each of the six positions varied from two (7%) to eight (27%).

The frequency of those responses corresponding to categories in each of the positions in the Treatment Videotape that was expected to occur by chance alone was compared to the data obtained. These expected frequencies were $1/6$ of the total number of subjects, $N=30$, or 5.0. That is, there were five chances out of 30 that the category of a subject's response would correspond to the category in a particular position of the Treatment Videotape,

assuming no seriation effect was functioning.

A chi square statistic was calculated, based on the expected and observed frequencies of responses corresponding to the six positions. This chi square was not significant at the .05 level ($\chi^2_5 = 4.8, p > .10$). Therefore, the null hypothesis that Initial Response Categories in all positions were equally likely to occur was not rejected. These results imply that the order of presentation of answers on the Treatment Videotape did not significantly affect the Initial Response Categories of subjects.

Representative Response Category

This analysis is similar to that above with the exception of the use of the Representative Response Category rather than the Initial Response Category as criterion for judging a response to correspond to a category of answer on the Treatment Videotape.

The distribution of the frequency with which subjects were judged to correspond to the category of answers in each of the six positions in the Treatment Videotape appears in Table 21.

TABLE 21
Distribution of Position in
Treatment Videotape of Represent-
ative Response Categories to the
Personal Question

| Position of Response Category | 1 | 2 | 3 | 4 | 5 | 6 | Total |
|-------------------------------------|---|---|---|---|---|---|-------|
| Frequency | 4 | 2 | 9 | 7 | 3 | 5 | 30 |

It can be seen from the table that the frequencies of responses do not differ markedly from those of Table 20.

The chi square statistic calculated with five degrees of freedom ($\chi^2_5 = 6.8, p > .10$), was not significant at the .05 level. The null hypothesis that Representative Response Categories in all positions were equally likely to occur was not rejected. This supports the conclusions of the previous

two analyses.

Taken together, the analyses that were used to investigate the hypothesis that the order of presentation of the model's "answers" would effect subjects' answers, did not support the contention that a seriation effect might be influencing the category of answers given by subjects. Subjects did not select a response from the Baseline Videotape or give a response after the Treatment Videotape any more frequently in a particular category because of the position of a corresponding answer in the videotaped sequence. There is no evidence to show that answers in all positions were not equally likely to be chosen or equated. However, a relatively greater number of matches occurred with responses in the third position of the Treatment Videotape. This is a result of the praise given for responses in these categories, as will be discussed in subsequent analyses.

Hypothesis 3

Hypothesis 3: The Category of Answer That Was Followed by Praise Will Occur More Frequently.

Having viewed a randomly assigned Treatment Videotape and Part III of Reflections on Time, subjects answered questions about the film. Effects of vicarious praise viewed on the videotape could be determined by comparing the number of times the category represented by a subject's response corresponded with the category of answer that was praised on the Treatment Videotape. If the praise was effective as a vicarious reinforcer, subjects' responses would match the praised category more often than expected by chance alone.

The subjects' responses to questions were classified as representing one of the six categories of the Coding Schedule: (i) Factualization; (ii) Comprehension; (iii) Understanding; (iv) Conjecture; (v) Application/Usage; or (vi) Evaluation. The effect of praise was analyzed by tabulating the number of times subjects' responses corresponded to (matched) the category of answer that was praised.

Hypothesis 3 was tested by comparing the frequency with which subjects' responses matched the category of answer that was praised with the frequency of matches that was expected to occur by chance alone. The following dependent variables were used to investigate the hypothesis: (i) the Initial Response Category to the Personal Question; (ii) the Representative Response Category to the Personal Question; and (iii) the Response Category to the Empathetic Question. A test of the hypothesis was carried out using each of these three variables.

Initial Response Category to the Personal Question¹

The frequency with which subjects were judged to match or not match the category of answer that was praised in the Treatment Videotape using the Initial Response Category as the criterion, is presented in Table 22.

¹ What would you answer to the question, "Can you tell me about the film?"

TABLE 22

Frequencies of Subjects Matching and Not Matching
the Praised Model Response with Initial Response
Category

| | Match | Not Match | Total |
|-----------|-------|--------------|-------|
| Frequency | 8 | 22 | 30 |

The eight matches indicated on Table 22 were spread over five categories of the Coding Schedule. Two matches with the praised category occurred when Factualization, Comprehension, and Conjecture "answers" were praised. The remaining two occurred when Understanding and Application/Usage responses were praised - one match for each category. No matches with the praised category occurred when that category was Evaluation.

To test Hypothesis 3, the frequency of observed matches (eight) was compared with the frequency of matches that was expected by chance alone. Since there were six treatments, corresponding to the six categories of answers, each of which was replicated five times, the overall probability of a match was one chance in six. That is, five subjects out of the 30 would be expected to match the praised category of response under conditions where no treatment effect was functioning. Using this chance probability of $1/6$, or $p=.1667$, binomial probabilities were calculated to determine the exact probability of a particular frequency of matches when $N=30$.² These probabilities are listed in Table 23.

2 Note that it was preferable to calculate exact probabilities rather than approximations using chi square analyses with one degree of freedom.

TABLE 23

Chance Probability of at Least X Matches
with Praised Model Response When $N=30$ and
 $p=.1667$

| Number of Matches (X) | Chance Probability of at Least X Matches |
|--------------------------|---|
| 0 | 1.00 |
| 1 | .97 |
| 2 | .90 |
| 3 | .76 |
| 4 | .58 |
| 5 | .38 |
| 6 | .22 |
| 7 | .11 |
| 8 | .05 |
| 9 | .02 |
| 10 | .006 |
| 11 | .002 |
| 12 | .0005 |
| 13 | .0001 |
| 14 | .00003 |
| 15 | .000005 |

It can be seen from Table 23 that the eight matches observed just reached significance at the .05 level. That is, eight matches would occur without a treatment effect in just 5% of the cases, so we can attribute the results to a treatment effect. Therefore, the null hypothesis that all categories of answers were equally likely to be matched was rejected. This implies that the praise of the model's answers was functioning as a vicarious reinforcer for the observing subjects. Subjects matched the praised category of response more often than was expected by chance alone.

Representative Response Category to the Personal Question

The frequency with which subjects were judged as matching the category of answer that was praised in the Treatment Videotape, using the Representative Response Category as criterion is presented in Table 24.

TABLE 24

Frequencies of Subjects Matching and Not Matching the Praised Model Response using Representative Response Category

| | Match | Not Match | Total |
|-----------|-------|-----------|-------|
| Frequency | 9 | 21 | 30 |

The nine matches, when using the Representative Response Category as criterion, presented themselves in the same categories as those occurring when using the Initial Response Category as criterion, with the one additional match being with a Comprehension "answer".

To test the hypothesis, the frequency of observed matches (nine in this case) was compared to the frequency of matches that was expected to occur by chance alone. Similar to the preceding analysis, the expected proportions of matches if no treatment effect was operating was $1/6$ of the total number of subjects. From the binomial probabilities as reported in Table 23, it can be seen that nine or more matches would have occurred by chance in just 2% of the cases when $N=30$ and $p=.1667$. Since this is below the 5% rejection level used in the present study, the null hypothesis that all categories of answers are equally likely to be matched was again rejected. Once again a treatment effect was evident. Subjects matched the praised category of response more often than was expected by chance alone.

Response Category to the Empathetic Question³

The frequency with which subjects matched the category of answer that was praised in the Treatment Videotape, when criterion for matching was the Response Category to the Empathetic Question was compared to the frequency of matches that was expected to occur by chance alone. The frequency of observed matches is found in Table 25.

TABLE 25

Frequency of Subjects Matching and Not Matching the Praised Model Response using Response Category to the Empathetic Question as Criterion

| | Match | Not Match | Total |
|-----------|-------|-----------|-------|
| Frequency | 5 | 8 | 13 |

Two of the observed matches with the praised category occurred when Comprehension-type answers were praised on the Treatment Videotape. The praise of Conjecture, Application/Usage, and Evaluation "answers" were also responsible for one match each. No matches occurred when Factualization and Understanding responses were praised.

Since only 13 subjects answered the Empathetic Question, the frequency of praised categories was not equal across all six categories. That is, different numbers of subjects watched the various categories of answers being praised. Therefore, $p=1/6$ was used only as a conservative approximate to the true chance probability of measures. Similar to the preceding analyses, to test the hypothesis, the frequency of observed matches (five) was compared to the expected frequency of matches, based on $N=13$ and $p=.1667$. Table 26 presents exact probabilities of obtaining particular numbers of matches.

³ What do you think (model), the girl on the videotape, would answer to the question, "Can you tell me about the film?"?

TABLE 26

Chance Probability of at Least X Matches with
Praised Model Response When N=13, $p=.1667$

| Number of Matches (X) | Chance Probability of at Least X Matches |
|--------------------------|---|
| 0 | 1.000 |
| 1 | .907 |
| 2 | .664 |
| 3 | .372 |
| 4 | .158 |
| 5 | .051 |
| 6 | .013 |
| 7 | .002 |
| 8 | .0003 |
| 9 | .00004 |
| 10 | .000003 |
| 11 | .0 |
| 12 | .0 |
| 13 | .0 |

It can be seen from Table 26 that the five matches just failed to reach significance ($p=.051$) at the .05 level. The null hypothesis cannot be rejected. However, the proportion of matches, 5/13, compares favourably with the proportion of matches observed in the preceding analyses:

| | |
|----------------------------------|------------------------|
| Initial Response Category | $\frac{8}{30} = 0.267$ |
| Representative Response Category | $\frac{9}{30} = 0.300$ |
| Empathetic Response Category | $\frac{5}{13} = 0.385$ |

The high proportion of matches suggest the possibility of a Type II error. That is, the proportion of matches here was greater than that of either of the preceding tests of Hypothesis 3; however, due to the small N the treatment effect was not

detected. Extrapolating from the power curve of $N=18$ (See Fig. 2, p. 76), when $p=0.385$, it appears that $1-\beta$ is less than 0.50. That is, there is less than a 50-50 chance of detecting a treatment effect with an N as small as 13 for this proportion of matches. Therefore, in spite of nonsignificance, it appears that these data are consistent with the treatment effect found in the preceding two analyses.

Discussion of Analyses testing Hypothesis 3

The first two analyses which investigated the existence of a treatment effect supported the hypotheses that vicarious praise strengthened answers of the desired category in subjects. Using both the Initial and Representative Response Category, a treatment effect was detected. The frequency of subjects' answers that were judged as matching the praised response category was above that which would be expected by chance alone. Although the final analysis failed to reach significance, on the strength of the relatively powerful tests using responses to the Personal Question as criteria, a treatment effect was confirmed.

The final and less powerful analysis dealing with Hypothesis 3 investigated the existence of a treatment effect in response to the Empathetic Question. This slight variation in the question from what the subjects themselves would answer (the Personal Question) to what they thought the model would answer (the Empathetic Question) created somewhat dramatic differences in responses from subjects. Few of the 30 subjects, only 13, would even attempt to respond, making the power of the test somewhat suspect. With such a small N , an effect may be present but not detected (Type II error).

Hypothesis 4

Hypothesis 4: The Category of Answer that Followed the Praise will not be Strengthened

As an alternative to the preceding hypothesis, there was a possibility that the vicarious praise had functioned as an attention-getter, thereby strengthening the category of answer which followed the praise rather than preceded it. Although it was thought that such an effect would not occur, the possibility needed to be explored.

The subjects' responses to questions were classified as representing one of the six categories of the Coding Schedule: (i) Factualization; (ii) Comprehension; (iii) Understanding; (iv) Conjecture; (v) Application/Usage; or (vi) Evaluation. The attention-getting effect of praise was analyzed by tabulating the number of times the subjects' responses corresponded to (matched) the category of answer that followed the praise.

Hypothesis 4 was tested by comparing the frequency with which subjects matched the category of answer which followed the praise with the frequency of such matches that was expected by chance alone. The Initial Response Category was used as criterion for matches because it provided one more match than the Representative Category and would, therefore, provide the strongest test of the hypothesis. If the results did not show significance when using the Initial Response Category as criterion they would not show significance when using the Representative Response Category.

The number of matches are indicated in Table 27.

TABLE 27
Frequency of Subjects Matching and Not Matching
the Category of Response Following the Praise,
using Initial Response Category to the Personal
Question as Criterion

| | Match | Not Match | Total |
|-----------|-------|--------------|-------|
| Frequency | 7 | 23 | 30 |

Two of the matches occurred with Factualization-type answers and with Comprehension-type answers. Understanding, Conjecture and Application/Usage answers were each responsible for one match. No matches occurred when an Evaluation answer followed the praise.

Since the 30 subjects were equally divided between groups that watched each of the six categories of answers following the praise, as before, the expected frequency of matches was $1/6$ of the total. According to the exact binomial probabilities for $N=30$, $p=.1667$ (See Table 23, p. 97), the seven matches did not reach significance at the .05 level. In fact, the probability of getting at least seven matches by chance alone is .11. Therefore, the null hypothesis that all categories of answers following the praise were equally likely to be matched by the subjects' answers was not rejected. There was no treatment effect demonstrating a strengthening of the category of answer that followed the praise. While not supporting the attention-getting function of the praise, the results did support the contention of the present investigation that a treatment effect would not be present to strengthen the category of answer which followed the praise.

Hypothesis 5

Hypothesis 5: More Intense Praise will have a Greater Facilitating Effect than Less Intense Praise

Since 30 was calculated to be an efficient sample size, additional volunteers experienced alterations of the original experimental treatments. One subsample experienced an alternative phrase of praise and another experienced alterations in the position of the praise on the Treatment Videotape. Data from the first of these cases will now be examined.

The praise used in the treatment sessions produced significant results when testing Hypothesis 3. Although a treatment effect was detected, it was not necessarily strong. Therefore, the proposed use of Treatment Videotapes with shorter forms of praise was abandoned in favour of the use of more intense praise to determine if a greater facilitating effect would result. This was consistent with the aim of the study which was the investigation of greater facilitation rather than less.

Hypothesis 5 was tested by comparing the frequencies with which subjects matched the praised category of answer when using the original form of praise and when using the more intense form of praise. Subjects' Initial Response Categories to the Personal Question were used in the analysis because this provided the strongest measure of an effect - more matches occurred with the Initial Response Category than with the Representative Response Category. If results did not show significance when using the Initial Response Category as criterion, they would not show significance when using the Representative Response Category.

Since the remaining 22 volunteer subjects were divided equally between this and the following test, 11 subjects were used to investigate Hypothesis 5. With this small number it was unreasonable to randomly assign subjects to six groups witnessing the praise of each of the six categories of responses. Therefore, only two categories of responses were used to investigate the alteration of praise. This was done in order to obtain similar group sizes to those which witnessed the original less intense praise. Moreover, in questioning the subjects who

witnessed the less intense praise, it was found that they generally classified the model's responses as being (i) simple and (ii) complex. Going on the assumption that subjects would be better able to differentiate between answers classified as "simple" and "complex" than between answers representing all six categories of the Coding Schedule, an attempt was made to show a treatment effect when more intense praise was given for just "simple" and "complex" answers. The "simple" and "complex" categories were represented by Factualization and Application/Usage answers, respectively.

The more intense praise for Factualization answers was "That's exactly the right kind of answer because you've remembered facts from the film." and for Application/Usage, "That's exactly the right kind of answer because you've shown how the film affected you." The less intense statements of praise were, "That's a very good answer. You've remembered what the film said.", and "That's a very good answer. You have applied what was presented.", for Factualization and Application/Usage, respectively.

The frequency of matches using more intense praise is indicated by Table 28.

TABLE 28
Frequency of Subjects Matching and Not Matching
the Praised Model Response When Employing More
Intense Praise

| | Match | Not Match | Total |
|-----------|-------|--------------|-------|
| Frequency | 3 | 8 | 11 |

Two of these matches occurred when the subject's response was judged to represent Factualization. The other match was with an Application/Usage response.

In analyzing the data, the expected probability of matching the praised category of answer was calculated from the number of matches occurring when the original phrase of praise was used. With the original praise, the proportion of matches with the Praised Model Response was 3/10 for Factualization and Application/

Usage responses. This proportion was used as the expected probability, $p=.3000$. The proportion of matching when using the original praise was compared to the proportion of matches observed with more intense praise, i.e. $3/11$ or $.272727$. Since the proportion of matches was less than when using the original praise, the difference between proportions of matches was in an opposite direction than hypothesized. The hypothesis that more intense praise will have a greater facilitating effect than less intense praise could not be confirmed.

The more intense praise, which was designed to be so by informing the model and subject that the praised answer was "exactly the right kind of answer," did not produce a stronger treatment effect. This could be attributed to the fact that subjects were already aware, without being specifically told, that the praised response was the right kind of answer when the less intense praise was employed. That is, more intense praise may have not added any additional information as to why the praise was given. As a consequence, the more intense praise was not acting as a stronger reinforcer.

Hypothesis 6

Hypothesis 6: Praise at the End of the Modelled Sequence will have a Greater Facilitating Effect than Praise in the Middle of the Sequence.

A second subsample experienced alterations in the position of the praise on the Treatment Videotape. By comparing the frequency with which subjects matched the praised category of answer when the praised answers came at the end and when they were interdispersed in the modelled sequence, as originally employed, it was possible to investigate the effects of the position of the P-reinforcer. Similar to the procedure with Hypothesis 5, the final 11 volunteer subjects were randomly assigned to a group watching either (i) Factualization or (ii) Application/Usage responses being praised. The treatment differed from the original in that the praised answers were the final two, rather than the third and ninth, given on the Treatment Videotape.

Subjects' Initial Response Categories to the Personal Question were again used in the analysis because they provided the strongest measure of an effect. The frequency of matches when praise was at the end of the modelled sequence is indicated by Table 29.

TABLE 29

Frequency of Subjects Matching and Not Matching the Praised Model Response When Employing Praise at the End of the Modelled Sequence

| | Match | Not Match | Total |
|-----------|-------|-----------|-------|
| Frequency | 1 | 10 | 11 |

The match, in this case, occurred when the subject's response was classified as representing the Factualization Category.

Expected frequencies of matches were calculated from data obtained from the first 30 codeable subjects. Therefore, the expected proportion of matches was $p=.3000$, based on the frequency of matches occurring for Factualization and Application/Usage answers with the original position of praise. The observed proportion of matches, $1/11$ or $p=.090909$ did not support the research hypothesis when compared to the expected proportion. Few matches occurred when the praise was at the end of the sequence. Contrary to the research hypothesis, subjects did not match the praised category of answer more often when those particular responses were at the end of the modelled sequence. The matching occurred because of some phenomenon other than a simple temporal relationship.

Hypothesis 7

Hypothesis 7: Aware Subjects are More Likely to Match Praised Model Response than Unaware Subjects

A subject was labelled as being "aware" or "unaware" of contingencies of reinforcement depending upon whether or not she was able to state the praise and give a description of the answer that was praised. If a subject gave those statements correctly it was assumed that she was aware of the general type of answer that gained praise. Furthermore, it was hypothesized that subjects being labelled as aware of the contingencies of reinforcement would be more likely to match the praised response than unaware subjects, since they would know what was required to gain the praise.

To test this hypothesis, the frequencies of aware subjects who did and did not match the Praised Model Response were compared to the frequencies of unaware subjects who did and did not match the Praised Model Response. The Representative Response Category of the first 30 codeable subjects was used in this comparison because it provided the strongest measure in favour of the aware subjects. If there was no significance using the Representative Response Category, there would be none using the Initial Response Category as criterion. The number of subjects in each of the two states (aware/unaware and match/not match) appear in Table 30.

TABLE 30

Awareness X Matching Praised Model Response
(Representative Response Category) for the
Personal Question

| | Match | Not | Total |
|-------|-------|-----|-------|
| Aware | 7 | 15 | 22 |
| Not | 2 | 6 | 8 |
| | 9 | 21 | 30 |

$p = .5473355$

It can be seen from the table that seven (32%) of the aware subjects and two (25%) of the unaware subjects matched the Praised Model Response.

Fisher's Exact Probability Test (Siegal, 1956, p. 96) was used to test the null hypothesis that awareness was unrelated to matching the praised response. Since it is an exact test, it is to be preferred over a chi square test corrected for continuity (Blalock, 1972, p.291). The exact probabilities were calculated by a computer programme written for that purpose. A listing of this programme appears in Appendix I. According to calculations based on the programme, the probability was .5473355 that observed frequencies of matches as extreme as in Table 30 would occur by chance. This was far from the .05 probability needed for rejection. Therefore, the null hypothesis that both aware and unaware subjects were equally likely to match the Praised Model Response could not be rejected.

An additional analysis was done to determine if subjects declared aware of the more intense praise matched the praised category more often than those declared aware of the original form of praise. Fisher's Test was again used to calculate the probability of the data set out in Table 31.

TABLE 31

Intensity of Praise of Aware Subjects X
Matching Praised Model Response (Representative
Category) for the Personal Question

| | Match | Not | Total | |
|--------------------------|-------|-----|-------|------------|
| Aware of Less Intense | 7 | 15 | 22 | p=.6307769 |
| Aware of More Intense | 2 | 5 | 7 | |
| | 9 | 20 | 29 | |

The chance probability of obtaining results as extreme as those in Table 31 was .6307769. Therefore, it was quite likely that the results were not related to awareness. This indicated that awareness of the more intense form of praise did not add to the experimental effect of matching the Praised Model Response. The present data support the discussion of results for Hypothesis 5 which suggested that the more intense praise did not increase awareness on the part of subjects. If awareness was increased an accompanying increase in matches would have been expected.

In fact, when Fisher's Test was used to calculate the probability of the observed frequencies of those who were labelled aware after watching either the less intense or more intense forms of praise, the results (See Table 32) showed $p=.4048925$.

TABLE 32
Intensity of Praise X Awareness for the Personal Question

| | Aware | Not | Total | |
|--------------|-------|-----|-------|--------------|
| Less Intense | 22 | 8 | 30 | $p=.4048925$ |
| More Intense | 7 | 4 | 11 | |
| | 29 | 12 | 41 | |

Frequencies as extreme as these observed could be expected to occur by chance about 40% of the time. Both those who watched the less intense and those who watched the more intense praise were equally likely to be declared aware of the contingencies. This also supports the contention, in the discussion of the results testing Hypothesis 5, that the subjects watching the less intense praise were already sufficiently aware of the contingencies in operation, thus not gaining any more information from the more intense praise.

Summary of Results Investigating Major Hypotheses

The first research hypothesis, that some categories of answers would be chosen more than others, was supported by the data. Some categories were represented more often than others when no praise was given for a particular category of answer.

Secondly, the research hypothesis that the order of presentation of the model's answers on the Baseline and Treatment Videotapes would effect the subjects' answers, was not supported. There was no seriation effect. Subjects' responses did not match the category of answer at the beginning or end of the videotaped sequences more often than that of answers in any other position.

The third research hypothesis, that the category of answer that was followed by praise would occur more frequently was supported; thus giving evidence in support of vicarious reinforcement of cognitive skills as exemplified by answering behaviours.

The fourth hypothesis, that the category of answer following the praise would not be strengthened was also confirmed. The praise did not act as an arouser to strengthen the type of answer which followed the praise. On the other hand, results from Hypothesis 3 showed that it did function as a vicarious reinforcer should, by strengthening the category of answer which preceded the praise.

The research hypotheses concerning the intensity and position of the praise were not supported. Neither a more intense form of praise nor praise at the end of the videotaped sequence had a greater facilitating effect than that of the original design.

Finally, the hypotheses that aware subjects would be affected more by the vicarious praise than unaware subjects was not supported by the data. Awareness was unrelated to whether subjects matched the praised response.

The analyses investigating the seven major hypotheses are summarized in Table 33 wherein null hypotheses, variables, and decisions concerning rejection are presented. The hypotheses investigated: (1) the frequency of all categories of answers when no praise is provided; (2) the influence of the ordering of the model's answers on those given by subjects; (3) the influence the vicarious praise on subjects' responses; (4) the influence of the vicarious praise on the category of answer which followed it; (5) the influence of more intense praise; (6) the influence of an alternate position of the praise; and (7) the influence of awareness of subjects' answers.

TABLE 33
Null Hypotheses, Variables and
Decisions used to Investigate
the Major Research Hypotheses.

| Null Hypothesis | Research Hypothesis | Independent Variable | Dependent Variable | Decision about Rejection |
|---|---------------------|--|--------------------|--------------------------|
| All categories are equally likely to occur by chance | 1 | Baseline Videotape | B.R.C. | Rejected |
| Response in all positions will be chosen an equal number of times by chance alone | 2 | Baseline Videotape | B.R.C. | Not Rejected |
| Categories of all positions are equally likely to be matched | 2 | Praised Model's Response (position) | I.R.C. | Not Rejected |
| Categories of all positions are equally likely to be matched | 2 | Praised Model's Response (position) | R.R.C. | Not Rejected |
| All categories of response are equally likely to be matched | 3 | Praised Model's Response (category) | I.R.C. | Rejected |
| All categories of response are equally likely to be matched | 3 | Praised Model's Response (category) | R.R.C. | Rejected |
| All categories of response are equally likely to be matched | 3 | Praised Model's Response (category) | R.C. | Not Rejected |
| The frequency of matching category <u>after</u> praise is at the chance level | 4 | Praised Model's Response (response that followed praise) | I.R.C. | Not Rejected |
| The frequencies of matching Praised Model Response was equal for less and more intense praise | 5 | Intensity | I.R.C. | Not Rejected |
| The position of the praise did not affect the amount of matching of Praised Model Response | 6 | Position | I.R.C. | Not Rejected |
| Aware and unaware subjects are equally likely to match Praised Model Response | 7 | Awareness and Praised Model Response | R.R.C. | Not Rejected |

B.R.C. = Baseline Response Category
I.R.C. = Initial Response Category
to Personal Question

R.R.C. = Representative Response
Category to Personal Question
R.C. = Response Category to
Empathetic Question

Minor Hypotheses

Several minor hypotheses were formulated, based on the information gained when investigating the preceding major hypotheses. First, alterations of categories were investigated. It seemed that the six categories of the Coding Schedule were not readily identified by subjects. However, subjects did classify the model's "answers" into broad categories which could be approximated by combining some of those in the Coding Schedule. Because of the developmental nature of the present study, a more efficient system of classification would serve to aid in the attempt to measure a subject's cognitive skill. Therefore, analyses based on the combination of categories were undertaken to determine if a stronger effect occurred using alternative classifications.

A second alternative coding of subjects' responses was undertaken in an attempt to investigate characteristics of those who answered the Empathetic Question.

Based on responses following the Treatment Videotape, subjects were identified as those who matched their baseline measure (coded 0), matched the praised category (coded 2), or matched some category not the same as their baseline or that which was praised (coded 1). It was thought that those coded 0, 1, and 2 may differ as to whether they answer the Empathetic Question.

The 0, 1, 2 coding could also be used to characterize a subject as being a "mover", those coded 1 or 2, or a "non-mover", those coded 0, and provided an alternative method of identifying subjects for additional analyses. It was thought that being a "mover" or "non-mover" was a learner characteristic which may somewhat predetermine a subject's response. Therefore, a comparison of "movers' " and "non-movers' " responses to the Personal and Empathetic Questions were investigated in an attempt to determine the extent of the "characteristic."

Third, in an attempt to investigate other learner characteristics, the consistency of a subject's response category to the Personal and Empathetic Question was investigated.

Consistency between response categories to these two questions would have given information about subjects' perceptions regarding the Personal and Empathetic Question.

The minor hypotheses which developed involved the following manipulations: (i) collapsing the categories of response according to those recognizable by subjects; (ii) determining if those whose responses matched their Baseline Response Category, matched the Praised Model Response, or matched some other response differed as to whether they answered the Empathetic Question; and (iii) determining if a subject's Response Category to the Empathetic Question was generally the same as the Response Category to the Personal Question. The rationale behind the minor hypotheses, statements of the hypotheses, and analyses of the data concerning each are now presented.

Collapsing Categories of Response

When questioned as to whether they noticed any difference in the types of "answers" given by the model, 23 of the 52 subjects were able to break down the "answers" into two or more types. The most common types mentioned were (i) facts, or simple answers; and (ii) the model's own ideas, or complex answers. The significant but far from substantial experimental effect of the vicarious praise found when subject's responses were classified according to the Coding Schedule, could have been due to the inability of subjects to sort out the complexities of model's responses based on the six categories of the Coding Schedule. Therefore, an examination of more fundamental categories, as suggested by the subjects themselves, was undertaken. In order to determine if the experimental effect (matching the Praised Model Response) was stronger for the classification suggested by the subjects or for the classifications of the Coding Schedule, the following collapsed categories were used to reclassify the responses:

- (i) Factualization, Comprehension, Understanding (1, 2, 3), and Conjecture, Application / Usage (4, 5)

- (ii) Factualization (1),
 Comprehension, Understanding (2,3) and
 Conjecture, Application/Usage (4,5)

The above regroupings of the categories were used because most subjects who were able to state different types of responses given by the model mentioned facts and opinions as two broad categories. Categories 1, 2, and 3 could be thought of as relating to facts drawn from the film, and categories 4 and 5 could be thought of as expounding opinion about the film. Furthermore, since categories 2 and 3 are related by definition more than category 1, they were separated out for analysis in the second regrouping. The sixth category was eliminated because few subjects' baseline measures and few responses to the Personal and Empathetic Questions were judged to represent the Evaluation category. Additionally, Evaluation (Category 6) was never identified by subjects as one of the categories of responses given by the model.

The original analysis was designed so that no subject observed an answer being praised that was of the same category as her baseline measure. Using the new collapsed schemes, however, it was possible that a subject's baseline measure might coincide with the category of answer which was praised on the Treatment Videotape. For example, using the (1,2,3) and (4, 5) classification, a subject with a Baseline Response Category identified as 4 and who watched Category 5 being praised on the Treatment Videotape, would watch a Treatment Videotape showing the praise of an "answer" of the same category as her baseline measure. That is, 4 and 5 are in the same category based on the collapsed scheme. Cases such as this were eliminated from analyses because a change in the category of response before and after treatment was desired. Moreover, the eliminations were necessary to attribute subjects' responses to the praise rather than to response predispositions.

The hypotheses that developed from the collapsed classification schemes were: each of the collapsed schemes will produce significantly more matches between subjects' response categories and the Praised Model Responses than will the original Coding Schedule.

For each of the collapsed schemes, the exact binomial probabilities were calculated. From the frequency of responses falling in each category in the baseline data and from the number of responses classified as matching the Praised Model Response using the original scheme, the expected number of matches using the collapsed scheme was calculated. The observed frequency of matches based on the collapsed schemes was compared to this expected number of matches.

Analysis based on the (1, 2, 3) and (4, 5) coding. For the (1, 2, 3) and (4, 5) scheme, there were 15 subjects because of the elimination of those who were identified as selecting a baseline response of the same category as that which was praised, and the elimination of those who watched Evaluation responses being praised. The Initial Response Category was used in the analysis because it provided the strongest measure of the effect. If an effect was not detected using the Initial Response Category it would not be detected using the Representative Response Category as criterion. Table 34 presents the frequency of matches with the praised response based on the (1, 2, 3) and (4, 5) coding.

TABLE 34

Number of Subjects Matching and
Not Matching the Praised Model Response
using a (1,2,3) and (4,5)
Coding

| | Match | Not Match | Total |
|-----------|-------|--------------|-------|
| Frequency | 10 | 5 | 15 |

Eight of the ten matches occurred when both the praised category and the subject's "answer" were classified as Factualization (Category 1). The remaining two occurred with Category 2 (Comprehension) responses. Therefore, all of the matches occurred when both the praised response and the subject's response were classified as belonging to the (1, 2, 3) collapsed category.

Based on the baseline and treatment data from preceding analyses, the chance probability of matches for the (1, 2, 3) and (4, 5) scheme was $p=.5216$. The binomial probabilities for $N=15$ and $p=.5216$ is found in Table 35.

TABLE 35
Binomial Probabilities for $N=15$, $p=.5216$

| Number of Matches (X) | Probability of at Least X Matches |
|--------------------------|--------------------------------------|
| 0 | 1.00 |
| 1 | .99 |
| 2 | .99 |
| 3 | .99 |
| 4 | .99 |
| 5 | .96 |
| 6 | .86 |
| 7 | .75 |
| 8 | .57 |
| 9 | .37 |
| 10 | .19 |
| 11 | .08 |
| 12 | .03 |
| 13 | .006 |
| 14 | .0008 |
| 15 | .00006 |

It can be seen from the table that the ten observed matches did not produce results significant at the .05 level. In fact, similar results would be expected to occur by chance in 19% of the cases. Therefore, the (1, 2, 3) and (4, 5) scheme did not detect the treatment effect. The collapsed scheme was not so efficient at detecting the treatment effect as the original Coding Schedule. One drawback here was that the low number of subjects, $N=15$, and

$p=.5216$ did not allow for a test as powerful as the test for $N=30$ and $p=.1667$ using the Coding Schedule. With $N=15$ and $p=.5216$ the effect may have been present but was not detected - a Type II error.

Analysis based on the (1) and (2,3) and (4,5) coding.

Using similar procedures as above, $N=19$ for the (1) and (2,3) and (4,5) coding, because of the elimination of those whose baseline measures coincided with the praised category of response and the elimination of those who watched Evaluation responses being praised. The number of observed matches was the same for both Initial and Representative Response Categories and is indicated on Table 36.

TABLE 36

Number of Subjects Matching and Not Matching the Praised Model Response using a (1) and (2,3) and (4,5) Coding

| | Match | Not Match | Total |
|-----------|-------|-----------|-------|
| Frequency | 8 | 11 | 19 |

Two of the matches occurred when both the praised category and the subject's answer were classified as Factualization (Category 1). The remaining six matches occurred when both were classified as either Comprehension or Understanding responses (Comprehension, Understanding - Coded (2,3))

The probability of matches, $p=.352$, was based on the Baseline data and the number of matches obtained from the original classification using the Coding Schedule. The binomial probabilities for $N=19$, $p=.352$ is found in Table 37.

TABLE 37
Binomial Probabilities for $N=19$, $p=.352$

| Number of Matches (X) | Probability of at Least X Matches |
|--------------------------|--------------------------------------|
| 0 | 1.00 |
| 1 | .99 |
| 2 | .99 |
| 3 | .98 |
| 4 | .94 |
| 5 | .85 |
| 6 | .70 |
| 7 | .53 |
| 8 | .34 |
| 9 | .19 |
| 10 | .09 |
| 11 | .04 |
| 12 | .01 |
| 13 | .003 |
| 14 | .0007 |
| 15 | .0001 |
| 16 | .00002 |
| 17 | .000001 |
| 18 | .00000006 |
| 19 | .0 |

From the table it can be seen that the chance probability of the eight observed matches is .34, far from the .05 rejection level. Therefore, the (1) and (2,3) and (4, 5) scheme did not detect the treatment effect. The collapsed scheme was not so efficient at detecting the treatment effect as the Coding Schedule. Again, it should be noted that with $N=19$ and $p=.352$, the test was not so powerful as when $N=30$ and $p=.1667$.

Summary. The analyses using the collapsed categories were undertaken to determine if the collapsed schemes were more sensitive at detecting a treatment effect than the original Coding Schedule. The subjects themselves classified the model's "answers" into categories similar to the collapsed categories which were used. If the original scheme was too complex for subjects to sort out, it was thought that the collapsed category schemes, being less complex, would be more sensitive to the experimental effect. The results did not show this to be so. The research hypothesis, that the collapsed schemes would be more efficient at detecting a treatment effect than the Coding Schedule, was not supported.

Matching Baseline Response Category, Praised Model Response, and Other Categories of Response as a Learner Characteristic.

In an attempt to investigate characteristics of subjects who responded to the Empathetic Question, data from the Personal Question were analyzed according to one final coding. It was possible to classify response categories according to those which were judged as matching the Baseline Response Category (coded 0), the Praised Model Response (coded 2), or some category other than the Baseline Response Category or Praised Model Response (coded 1). The subjects coded 0 were those who did not change from their Baseline Response Category when responding to the Personal Question. Those coded 1 and 2 did change, with 2's changing to that category which was praised and 1's changing to something other than that which was praised.

0's, 1's and 2's answering the Empathetic Question. With so few subjects answering the Empathetic Question, i.e. N=13, it was possible that making responses to it was dependent upon some learner characteristic. Identifying subjects as staying with their baseline measures (0's), moving to the praised category (2's), or some other category (1's), when responding to the Personal Question, analyses were done to determine if being a 0, 1, or 2 was a predominate characteristic of those who responded to the Empathetic Question.

An attempt was made to determine if those who were coded 0, 1, or 2 (based on the Representative Response Category since it gave the strongest measure of the vicarious praise) differed as to whether they responded to the Empathetic Question.

The frequencies of subjects coded 0, 1, and 2, answering and not answering the Empathetic Question are found in Table 38.

TABLE 38
Code 0, 1, and 2 X Answering the Empathetic Question

| Code | 0 | 1 | 2 | Total |
|---------------|---|----|---|-------|
| Answering | 4 | 3 | 6 | 13 |
| Not Answering | 5 | 9 | 3 | 13 |
| | 9 | 12 | 9 | 30 |

Even without correcting for continuity, the chi square calculated from this data was not significant at the .05 level ($\chi^2_{.5}=3.6427$, $p > .10$). Therefore, the null hypothesis that 0's, 1's and 2's were equally likely to answer the Empathetic Question was not rejected. Answering the Empathetic Question was unrelated to prior coding as a 0, 1, or 2.

"Movers" and "Non-movers" answering the Empathetic Question.

Another way of investigating the question as to whether 0's, 1's and 2's (based on the Representative Response Category) differed in their frequency of responding to the Empathetic Question was to consider "movers" and non-movers". A "mover" was identified as a subject who changed from the Baseline Response Category to some other category (those coded 1 or 2) and a "non-mover" was identified as a subject who stayed with the Baseline Category (those coded 0). It was thought that perhaps "movers" and "non-movers" differed as to whether or not they answered the Empathetic Question. The frequency of "movers" and "non-movers" answering and not answering the question is found in Table 39.

TABLE 39
 Mover vs. Non-mover X Answering
 the Empathetic Question

| Code | Mover | Non-mover | Total |
|---------------|-------|-----------|-------|
| Answering | 4 | 9 | 13 |
| Not Answering | 5 | 12 | 17 |
| | 9 | 21 | 20 |

It can be seen that the "movers" accounted for four (31%) of those who answered the Empathetic Question, and five (29%) of those who did not, while "non-movers" accounted for nine (69%) of those who answered, and 12 (71%) of those who did not.

The exact probability of chance results as extreme as those shown in the contingency table was calculated using Fisher's Test (Siegal, 1956, p. 96). The programme for Fisher's Test is shown in Appendix I. The probability of $p=.6227886$ indicated there was no relation between a subject's answering the Empathetic Question and being a "mover" or "non-mover". The likelihood of the observed results occurring by chance was far greater than the .05 rejection level. Therefore, the null hypothesis that "movers and "non-movers" were equally likely to answer the Empathetic Question was not rejected.

"Movers" answering the Empathetic Question. An analysis was also undertaken to determine if those coded 1 and 2 differed as to whether they answered the Empathetic Question. It was thought that matching the Praised Model Response, coded 2, might increase the probability of responding to the Empathetic Question when considering only those labelled as "movers". The frequency of 1's and 2's answering and not answering the Empathetic Question is found in Table 40.

TABLE 40

Code 1 or 2 X Answering the Empathetic Question

| Code | 1 | 2 | Total |
|---------------|----|---|-------|
| Answering | 3 | 6 | 9 |
| Not Answering | 9 | 3 | 12 |
| | 12 | 9 | 21 |

The frequency of 1's who answered the Empathetic Question was three (33% of those answering), while the frequency of 2's answering was six (67% of those answering). The exact probability of the observed data, using Fisher's Test, was calculated as $p=.07132651$, approaching but not reaching the .05 rejection level. Therefore, the null hypothesis the 1's and 2's did not differ as to whether or not they responded to the Empathetic Question could not be rejected. On the whole, it appeared that staying with the baseline or moving to either the praised response category or some other response category was not characteristic of those who answered the Empathetic Question.

Answering the Personal and Empathetic Questions with the Same Category of Answer

In an attempt to identify consistency as a characteristic of subjects who answered the Empathetic Question, an analysis was done to determine if the frequency of subjects judged as answering both the Personal and Empathetic Questions with the same category of response differed from that which would be expected by chance alone. For the Personal Question, the Representative Response Category was used because it provided the strongest measure, that is, a greater number of matches than with the Initial Response Category. If the results were not significant using the Representative Response Category, they would not be significant using the Initial Response Category as the criterion. The number of subjects who were judged as responding to both questions with the same category of answer is found in Table 41.

TABLE 41

Frequency of Subjects Responding with Same and Different Response Category for Personal Question (Representative Response Category) and the Empathetic Question

| | Response Category Same | Response Category Different | Total |
|-----------|------------------------------|-----------------------------------|-------|
| Frequency | 4 | 9 | 13 |

It can be seen that four subjects (31% of those answering the Empathetic Question) were classified as answering with the same response category as their response category to the Personal Question.

A conservative estimate of the expected frequency of matches was used in the chi square analysis. That is, an expected frequency of $1/6$ of the total was used for expected frequency of subjects answering both questions with the same category of answer. From the binomial expansion for $N=13$, $p=.1667$ (See Table 26, p.100) it can be seen that the four matches were far from significant at the .05 level ($p=.158$). Therefore, the null hypothesis that the frequency of matches between the Category of Response for the Personal and Empathetic Questions was at the chance level, was not rejected. Subjects who were labelled as answering the Personal Question with a particular category of response did not answer the Empathetic Question with the same category of response above the chance level. This did not support the research hypothesis that those who answered the Empathetic Question would have answered with the same category of response as their answer to the Personal Question.

Summary of Results of Minor Hypotheses

Results dealing with the minor hypotheses indicated that collapsed categories were no better at detecting a treatment effect than were the six categories of the Coding Schedule. Although Type II errors may have occurred, the simpler classifications schemes were not to be preferred over the Coding Schedule considering the present analyses. Moreover, subjects coded 0, 1, and 2 did not differ as to their frequencies of responding to the Empathetic Question; and, finally, the frequency of those who answered both questions with the same response category was at the chance level. No characteristics of subjects who answered the Empathetic Question were identified. Responding to the Empathetic Question was independent of the characteristics investigated, i.e. a 0, 1, 2, coding and consistency of answers. The null hypotheses, variables, and decisions for each of the minor hypotheses are presented in Table 42.

TABLE 42

Null Hypotheses, Variables and Decisions
Used to Investigate the Minor Hypotheses

| Null Hypothesis | Independent Variable | Dependent Variable | Decision About Rejection |
|---|---|--|--------------------------|
| All categories are equally likely to be matched for a (1,2,3) and (4,5) coding | Praised Model Response according to (1,2,3) and (4,5) | Initial Response Category to Personal Question | Not Rejected |
| All categories are equally likely to be matched for a (1) and (2,3) and (4,5) coding | Praised Model Response according to (1) & (2,3) & (4,5) | Initial Response to Personal Question | Not Rejected |
| 0's, 1's and 2's were equally likely to answer Empathetic Question | Matched Baseline (0) Praised (2) or other (1) based on R. R. C. | Answering Empathetic Question | Not Rejected |
| "Movers" and "non-movers" were equally likely to answer Empathetic Question | Matched Baseline (0) or Praised or other (1 and 2) | Answering Empathetic Question | Not Rejected |
| 1's and 2's were equally likely to answer Empathetic Question | Matched Praised (2) or other (1) | Answering Empathetic Question | Not Rejected |
| Frequency of those who answered same category for Personal Question and Empathetic Question was at the chance level | R.R.C. for Personal Question | Response Category for Empathetic Question | Not Rejected |

R.R.C. Representative Response Category

CHAPTER FIVE

Discussion and Conclusions

This chapter presents a discussion of: (i) the results of the present study; (ii) experimental, educational and general implications of those results; (iii) limitations of the investigation; and (iv) recommendations for further research. The final summary provides a concise account of the study as it developed and the results thereof.

Discussion of Results

In normal classroom situations many variables influence the amount of learning in any given instance. Within the controlled environment contrived for the present study, it was possible to partition and manipulate a few variables in order to determine their effect on behaviour. The behaviour in this case was cognitive skill as manifested by specific types of answers. Although many theories may exist to explain answering behaviour in classrooms, the precisely defined learning situation of the present study reduced the variables which could be held responsible for the answers given by subjects. Within the controlled situation, differential effects of characteristics of the stimuli were investigated by the following questions:

- (i) Were subjects' responses influenced by the order of those observed?
- (ii) Were subjects' responses influenced by the praise they observed, and which types of responses were strengthened?
- (iii) Were subjects' responses influenced by the intensity and position of the praise they observed?

A discussion of how the characteristics of the stimuli mentioned in the above questions related to the answering behaviour of subjects follows. Also, since characteristics of the criteria for classifying responses may have influenced the results of the present investigation, alternate classifications are explored.

Conclusions drawn from these alternate classification schemes are discussed.

Stimulus Characteristics Influencing Learning

Influence of order. Although theory suggested that the order of the modelling stimuli would influence the responses subsequently given by subjects (See for example Hovland, 1938), the order of responses given by the model was not a crucial factor influencing the subjects' responses. The hypothesis that categories of responses in all positions of the model's script were equally likely to be matched by subjects could not be rejected. Specific categories of response seemed to be more influential than the position of those categories. That is, subjects' response categories were influenced by the types of responses given by the model rather than by the position of the responses in the stimulus material.

Verification came from the baseline measures taken on subjects. The order of the model's "answers" on the Baseline Videotape did not influence subjects' baseline measures, but particular characteristics of the model's "answers" did. Subjects selected baseline responses over all six of the categories of the Coding Schedule, regardless of their position in the Baseline Videotape, some categories being chosen more often than others.

Influence of praise. In keeping with the reinforcement theory put forth by Bandura (1969) and others (Miller and Dollard, 1945), the results of the present study showed that vicarious praise influenced answering behaviour of observing subjects. Subjects increased desired categories of answers (to the Personal Question) when "answers" representing those categories preceded praise on the Treatment Videotape. Contrary to the hypothesis that praise would serve as an attention-getter and, thereby, strengthen the category of response following rather than preceding it, the praise did function as a vicarious reinforcer. In fact, a strengthening of the response category following praise was highly unlikely since, by design of the experiment, the praise was directed specifically to the category of response preceding praise.

Using Bandura's mediational theory to explain the detected effect, subjects were (1) attentive to the modelling stimuli, (2) able to store the information gained from watching the model, (3) capable of reproducing similar answers, and (4) motivated by the praise given to the model. Subjects were able to code the information they gained from the Treatment Videotape via internal symbolic control. They "identified" with the model sufficiently to realize, sometimes at a nonverbal level, that if she was praised for a certain category of answer, they too might be praised for a similar type of answer. Skinner (1974) would explain the same learning by stating that observers responded "indirectly" to praise on the videotape and were subsequently rewarded for matching the praised category of response. In this case, Bandura's explanation appears more valid since learners were not overtly rewarded for matching.

It should be noted that subjects were required to generalize from the responses given by the model to the type of responses required of themselves. That is, since the model and subjects were questioned about different material, subjects had to generalize from the one specific type of praised "answer" to new material. Specific response characteristics were required from the subjects before they could be classified as matching the model's praised response. It is quite possible that some subjects were not classified as matching the Praised Model Response because they were attentive to characteristics that were not included in the criteria for matching. Nevertheless, a sufficient number of subjects were able to generalize from the modelled sequence to produce significant results related to the Personal Question.

Although the praise functioned as a vicarious reinforcer when analyzing subjects' answers to one question (the Personal Question) the effects were not so evident when analyzing subjects' answers to a second question (the Empathetic Question). This second question required subjects to answer for the model rather than for themselves. The vicarious praise did not significantly affect answers to the Empathetic Question, but the proportion of matches was consistent with results from the Personal Question. Therefore, the insignificance

may have been due to the low power of the test and accompanying difficulties in detecting effects. In spite of the fact that data from the Empathetic Question failed to show a significant experimental effect, the notion of vicarious reinforcement of cognitive skill was not totally rejected.

Nonetheless, it should be noted that the difference between results from the two questions could have resulted from a number of reasons not investigated in the present study. For instance, the differences may have occurred because of subjects' inability to identify with the model. That is, the wording of the Empathetic Question may have made subjects unwilling to answer, since answering for the model implied a closer identification with her than was needed when responding to the Personal Question. Alternately, subjects may not have perceived the praise as potent enough to influence the model - they were not aware that it affected their own answers in many cases. No conclusions regarding the effects of identification with the model, perception of the praise, or similar characteristics of the learners could be drawn from data collected in the present investigation. However, a comparison of responses to both the Personal and Empathetic Questions did imply that subjects interpreted the questions as being entirely different although the wording was very similar.

What was found in the present investigation was the first manifestation of vicarious learning. According to Bandura, a behaviour can be learned vicariously without being performed, simply by observing another's behaviour. However, both Skinner and Bandura agree that for long term effects, behaviour must be strengthened by reinforcers given direct to the performer. Therefore, in order to assure a strong effect in the future, those who learned vicariously would require incentives in the form of potential reinforcers.

Influence of position and intensity of praise. After the initial experimental data demonstrated vicarious reinforcement of cognitive skill, certain stimuli were altered in order to explore their relative effects.

First, it was thought that more intense praise would have a greater facilitating effect. However, the more intense praise did not produce a greater treatment effect in the present work.

Although Bandura (1969) suggests that qualities of the reinforcing stimuli are influential in regulating the amount of vicarious learning and Zahorik(1969) states that rarely used (and therefore more intense) forms of feedback are better facilitators of behaviours, the expected increase of desired responses did not occur. The less intense praise must have already indicated to subjects that the praised response was desired.

Additionally, the change in position of the praise did not produce greater treatment effects, contrary to what was expected. Because of a possible seriation effect (See for example Hovland, 1938) as mentioned earlier, and Bandura's (1969) contention that characteristics of the reinforcing stimuli affect the potency of reinforcers, a greater facilitation of the desired category of response was expected when the praise was moved to the end of the modelled sequence. However, it seems that the praise was potent enough that a change in position did not alter the amount of attention given to it and did not alter its reinforcing qualities. This is in agreement with the analyses investigating a seriation effect. Subjects were influenced by response categories rather than the position of responses on the Treatment Videotape. The number of responses given by the model was not so great that the subjects forgot the category of response that was praised when the praise was not at the end of the sequence.

Alternate classifications. In addition to using the Coding Schedule, two alternate methods of classifying the data were examined in an attempt to demonstrate other ways of detecting a treatment effect.

It was hypothesized that more matches would occur with collapsed categories when compared to the number occurring with the six categories of the Coding Schedule. Since the subjects would be better able to sort out the modelling stimuli if fewer categories were represented and if categories were similar to those identified by subjects, it seemed that a greater treatment effect should appear

when either of two collapsed schemes were employed. However, the initial method of classifying subjects as matching or not matching the praised category of response according to the Coding Schedule, was found to be more efficient at detecting the treatment effect than the alternate methods. The alternate methods of classifying the data did not aid in illuminating the treatment effect.

These results may have occurred because an amount of data was lost when the categories were combined. That is, with collapsed categories, some subjects were eliminated from analysis because their baseline measure corresponded to the praised category or because they had observed an unpopular category of answer (Evaluation) being praised. A smaller N made it more difficult to detect a treatment effect if it was present (Type II error). While the results did not imply superiority in classifying with the collapsed schemes, the possibility of Type II error suggests further investigation of alternative schemes.

Learner Characteristics Influencing Learning

Besides controlling the stimulus materials and the experimental environment, it was possible to investigate a number of learner "characteristics" to determine their effects on the amount of learning. The following questions dealing with learner characteristics were investigated:

- (i) Was awareness of contingencies of reinforcement a factor which influenced the amount of vicarious learning?
- (ii) Was movement from a baseline measure characteristic of subjects answering the Empathetic Question?
- (iii) Was consistency of responses for both the Personal and Empathetic Questions characteristic of subjects?

These characteristics will now be discussed according to how they influenced the amount of learning.

Awareness. Awareness of contingencies of reinforcement was not related to the amount of vicarious reinforcement. That is, the desired category of response was sometimes strengthened without the subject being able to state the praise and the type of response which had been praised. In agreement with research

such as that by Binder, McConnell, and Sjöhom, 1957; Ditrachs, Simon and Greene, 1967; Greenspoon, 1955; Krasner, Weiss, and Ullman, 1959; and Ryan and Krumboltz, 1964, the present study lended support to the contention of learning without awareness.

Results from the Empathetic Question also tend to discredit awareness as an influential characteristic. That is, subjects chose not to answer the Empathetic Question by stating that they had no idea as to how the model would answer the question "Can you tell me about the film?". If awareness influenced the vicarious reinforcement, subjects who were aware of the contingencies of reinforcement would have been expected to respond to the Empathetic Question with a response similar to that which was praised. Although 62% of those who answered the Empathetic Question were labelled as being aware, only 38% of that 62%, or about 24% of the aggregate, matched the Praised Model Response. Moreover, 59% of those who were declared aware did not answer the Empathetic Question. Therefore, in the present study, awareness as defined herein appeared to be unrelated to answering the Empathetic Question.

Nonetheless, the other side of the debate proposes that there can not be learning without awareness. See for example Dulany, 1961; Kriekhouse and Eriksen, 1960; Levin, 1961; Matarazzo, Saslow, and Pareis, 1960; Spielberg and DeNike, 1962; and Tatz, 1960. These studies suggest that results showing learning without awareness are attributable to an inability to detect the awareness because of insensitive instruments. This could be true of the present study, although every attempt was made to give opportunity to subjects to state awareness of contingencies. The learning without awareness result could stem from the fact that the modelled stimuli were too complex for the subjects to verbalize all they observed; they were therefore unable to state the phenomena which affected them. It appears that fewer categories of answers given by the model may alter the results. With fewer categories, the subjects would have a greater chance of discriminating between the types of answers that were praised and were not; that is, they would be cognizant of the critical response category more quickly.

Movement from baseline. The data showed that answering the Empathetic Question was independent of whether the subject changed from her Baseline Response Category when answering the first question. Staying with the Baseline Response Category, changing to the praised response category, or changing to any other response category for the first question did not influence the answering of the second. Prediction concerning the answering of the Empathetic Question could not be made on the basis of a subjects' predisposition to baseline measures.

Consistency. On an empirical basis, it was expected that those who matched their Baseline Response Category, the praised category, or some other category when answering the Personal Question would also do so when answering the Empathetic Question. Although the modelling stimuli had been constructed to provide as much similarity as possible between the model and subjects, the expected results did not occur. Bandura (1969) states that "identification" with the model is important in determining the amount of vicarious learning. In this same vein, the lack of expected results could stem from the fact that subjects did not readily identify with the model. Other reasons not investigated in the present study, could also contribute to the results. For instance, subjects could have histories of being rewarded for giving different types of answers to similar questions asked adjacent to one another (the questions under investigation here were similar and adjacent to one another), since there is an increased probability that one of the two answers is correct and could be rewarded.

Going on the assumption that subjects interpreted the Personal Question (responding for themselves) and the Empathetic Question (responding for the model) as being similar, it was expected that their response categories would be similar for both questions. This was not so. Once again it appeared that subjects were not influenced by the similarity between the questions or by the similarity between themselves and the model. It appeared that subjects perceived the two questions as being completely different and responded to them in that light.

Past reinforcement histories were perhaps more influential in that they may have been rewarded for giving different types of answers to similar questions.

Summarizing the findings concerning learner characteristics; (i) awareness of contingencies did not influence the amount of vicarious learning; (ii) movement from the baseline measure was not characteristic of those who answered the Empathetic Question; (iii) consistency of responses was not characteristic of those who answered the Empathetic Question. If specific learner characteristics were affecting the amount of learning and the occurrence of answers to the Empathetic Question, they were not the ones examined.

One of the limitations of reinforcement theory is the difficulty in investigating all variables which may theoretically influence learning. For instance, answering questions similar to the Empathetic Question, i.e. answering for others, may not have been previously rewarded. However, answer questions similar to the Personal Question probably have been rewarded previously. Such histories may influence the results of the present study.

Experimental Implications

Stimulus and learner characteristics manipulated in the present investigation have certain broader experimental implications. Some of these implications were derived from the developmental nature of the present investigation. Flexibility in the developmental stages allowed modifications of the materials and experimental procedures. These modifications included alterations of (i) the forms of question used; (ii) the age of subjects; and (iii) the experimental design.

The initial short-answer completion and multiple-choice items elicited many types of responses from subjects. However, because it was desirable to obtain some measure of spontaneous answering behaviour in order to attribute learning to the praise rather than to the specific question asked, open-ended questions were subsequently employed. This precluded the study of direct reinforcement and the direction turned toward vicarious reinforcement.

Nevertheless, it may be possible to use a greater variety of item forms than those employed in the present experiment. To determine the effect of praise on answering behaviour, subjects could be asked short-answer or multiple-choice questions representing all types of cognitive skills for a baseline measure. Then praise could be given for answers to certain types of questions. A final measure of the treatment effect could be obtained by asking open-ended questions. This type of questioning may, in fact, be more related to the normal classroom situation where both short-answer, multiple-choice, and open-ended questions are asked. Vicarious learning of the above situation could be tested by questioning the subjects who watched the praise of others.

Also related to the questions employed in the present work, it appeared that subjects reacted differently to the Empathetic and Personal Questions, i.e. fewer responded to the Empathetic Question. A subject's affiliation with the model and awareness of contingencies of reinforcement may be influencing this result. As previously discussed, improved measures of awareness may show awareness of contingencies to be characteristic of those who answer questions such as those as the Empathetic Question. Also, model characteristics need further exploration to determine salient factors which would influence answers to questions similar to Empathetic Questions.

Initially there was flexibility in the age of subjects used in the present study. Older subjects were sometimes found to answer high level answers at the expense of lower level answers and were, therefore, unsuitable for the purposes intended. However, if cognitive skill was manifested by some behaviour which would not show this ceiling effect, older subjects could suitably be used in similar experimental situations.

Moreover, the developmental work permitted the flexibility to design an experimental situation for the examination of vicarious reinforcement. Since most vicarious reinforcement studies have dealt with classroom maintenance behaviour, there were few guidelines for the construction of a situation that would allow for the examination of cognitive skill as manifested by answering behaviour. The experimental situation itself was so arranged as to attribute a

substantial amount of vicarious learning to the praise on the videotape. This, of course, goes on the assumption that all other variables were controlled so that any strengthening of the desired answering behaviour could be attributed to the vicarious praise. If other independent variables could be shown to influence the learning, then they would also be identified as vicarious reinforcers. A videotaped model in a controlled situation similar to that used in the present investigation could be used to test the effectiveness of these other P-reinforcers.

In the present study, the small number of subjects used to test the effects of varied intensity and position of praise limited the conclusions which could be drawn. Therefore, additional work is needed to investigate salient characteristics of the stimuli observed by subjects. Additionally, the strengthening of answering behaviour as a manifestation of cognitive skill implies that other manifestations, precisely defined by overt behaviours, may also be affected by vicarious reinforcers.

Furthermore, the experience with the Taxonomy in the present study implies that researchers should carefully investigate the Taxonomy in light of their own research problem to determine its suitability.

For the present investigation, a modification of the Taxonomy was found to be more satisfactory than the Taxonomy itself. Although this does not totally discredit the use of the Taxonomy, it does imply that indiscriminant use of it may be undesirable. The Taxonomy appears to be useful for classifying objectives according to their cognitive level, although the levels may not be strictly hierarchical. Answers corresponding to questions for specific objectives may not necessarily represent a true hierarchy of answers. The classification of answers is too subjective and dependent upon individual past histories to validly use the Taxonomy for their classification.

Educational Implications

Although the experimental situation was somewhat removed from what is generally found in classrooms and the sample was limited to teenage girls, a certain amount of useful information was obtained.

First, if teachers aim to increase cognitive skills as an educational goal, it is helpful to know some of the stimuli which can potentially influence those skills. The present study showed answering behaviour as it was influenced by vicarious praise. Although different manifestations of cognitive skill may be under scrutiny by teachers, the present study implies that they might attempt to define the reinforcers which can be used to strengthen the desired skill. The fact that praise was an effective reinforcer in the present study served merely to confirm previous research and teachers' experiences.

Moreover, the present work implies that when teachers struggle to elicit specific responses from students, they might well try "modelling" as a technique to hasten the operation. Students can learn a great variety of behaviour vicariously. These behaviours range from attending to work to answering questions in specific ways. Within the modelling framework, the specific characteristics of the observer and model students would need to be examined to determine effective models for specific students.

In keeping with the aim of the present investigation to determine if cognitive skill could be shown to be influenced by vicarious reinforcers, the experimental setting allowed the necessary controlled situation but took a step away from the normal classroom situation. Thus, although strict control was maintained, the situation may not be entirely generalizable to the normal classroom situation. The teacher does not have control over the entire spectrum of reinforcers available to students in classrooms. For this reason teachers should look at the significant results and remember that such significance may not necessarily be evident in classrooms. However, if students are told the types of cognitive skills and these skills are demonstrated by "models" even more efficient learning may occur than that illustrated in the present study.

General Implications

The results of the present investigation imply that, in general: (i) vicarious reinforcers control a much broader spectrum of behaviour than research shows; (ii) operant conditioning is a slow process that can be hastened by modelling; and (iii) classroom verbal interaction may be influenced by vicarious reinforcers. These will be discussed in that order.

In an attempt to add to the general research concerning vicarious reinforcement, the present study focused on cognitive skill as manifested by answering behaviour. Research has also shown questioning behaviour to be influenced by vicarious reinforcement. (See for example Krumboltz and Schroeder, 1965). It appears that a greater variety of cognitive skill than has been investigated is under the control of vicarious reinforcers. Additionally, cognitive skills that are readily manifested may also be under the control of direct reinforcers. Although the answering behaviour of the present study occurred too infrequently to investigate the operant control of such behaviour, others have shown cognitive skills to be influenced so (See for example Goetz and Baer, 1973).

However, it should be noted that operant conditioning techniques, i.e. direct reinforcement of successive approximations to the target behaviour, is often an inefficient method of modifying behaviour. As was found in the present study, a model can provide consistent clues which can assist the learner when desired behaviour has low frequencies.

Besides relating to research and theory dealing with vicarious reinforcement, the present investigation is related to the theory of classroom verbal interactions. As Nuthall (1971) points out, students most often listen to a question-answer exchange between the teacher and some other student; and the covert question-answering done by the observer is the prime determiner of the learning that takes place in the observer.

Nuthall (1971) suggests that, because of feedback to the participant being observed, the observer is likely, if given the opportunity to speak, to respond with an answer which is "thought to have the correct content and form (p. 12)". Additionally,

he suggests that some of the complexities of information available to the observers include: (1) the precision of the information available to the observers - is it clear what responses are acceptable as correct and which are not; and (2) the length of the questioning leading up to information about the correct response - what is the proportion of correct responses. These complexities are similar to those discussed by Bandura (1969).

Although covert question-answering as put forth by Nuthall (1971) is difficult to measure, the present study did illustrate manifestations of learning from observing, supporting his theory, by operationalizing cognitive skill as overt answering behaviours. In spite of the complexities as suggested above, the present investigation was able to illustrate the operation of vicarious reinforcers in the strengthening of cognitive skill. Nuthall (1971) also alludes to the operation of vicarious reinforcement when he states, "If the observed response matches the observing subject's own, then the teacher reaction has a direct reinforcing effect on the processes involved in constructing the covert response, (p. 20)".

Limitations of the Study

From the results of the present investigation, it was concluded that vicarious praise does strengthen specified answering behaviours, illustrating the vicarious reinforcement of cognitive skill. The order of the model's responses and changes in intensity and position of the vicarious praise did not significantly affect the amount of vicarious learning and it was not necessary for subjects to be aware (as measured presently) of the contingencies of reinforcement for them to be influenced by the praise. Although these conclusions could be drawn from the results, certain limitations of the study govern the generalizability of the results.

These limitations will be discussed in the following order: (i) the experimental setting; (ii) the sample used; (iii) the size of subsamples; (iv) the classification scheme, (v) the representation of cognitive skill; and (vi) the nature of the data.

Experimental setting. Although Krumboltz and Thoresen (1964) found individual and group settings equally effective when exploring learning through direct and vicarious reinforcement, and although it is sometimes necessary to compartmentalize variables for close examination (Skinner, 1974), it is often desirable to study effects under conditions more representative of those to which the study will be applied (See Snow, 1974). The restrictedness of the setting of the present study somewhat limits its generalizability (the results may not be typical of all classroom situations) although it was possible to control certain variables in order to obtain a close investigation of vicarious praise.

Limited sample. In order to control for unwanted variance which could occur by using subjects of both sexes and in order to have as much similarity as possible between the model and the subjects, both a female model and female subjects were used. Thus, the data may be representative of subjects watching a same-sex model but this cannot be determined until similar data from male subjects watching a male model are obtained. Additionally, the data suggest that the effect would be functioning for both male and female subjects regardless of the sex of the model but further work would be needed to show this explicitly since the nature of the task may be sex-related and may, therefore, alter the amount of learning occurring with males as compared to females.

The age of the subjects, 15 to 16 years, may also limit the generalizability of the data. Students of different ages may react differently to the modelling situation as represented in the present study.

Small subsample size. Although the sample size of 30 was calculated to be reasonably sufficient for detecting an experimental effect, the use of additional volunteer subjects in alterations of the original design resulted in interesting but weak tests. For instance, the change in intensity of the praise did not produce significant results (more intense praise did not produce more matches with the Praised Model Response than less intense praise). The null hypothesis that both types of praise (more and less intense) were equally effective could not be rejected based on the data.

However, the results may have occurred because the effect existed but could not be detected (Type II error) because of the small N. Increasing the subsample to investigate the increased intensity of praise would result in a more powerful test and more convincing conclusions. Larger samples are also needed to investigate more fully the change in position of the praise and alternate classification schemes.

The use of additional volunteer subjects in alterations of the original design permitted efficient use of volunteers since only 30 were needed for the original treatments. However, the small subsamples used in alterations limited the power of the tests and, therefore, the results.

Categories of answers. Although the Coding Schedule appeared to be ideal for use in the present investigation (producing more consistent coding than the Taxonomy), it may have limited the study to some extent. For instance, none of the subjects in the present study were able to state all six types of answers given by the model. This may have affected their chances of being declared aware of the contingencies of reinforcement - being unable to state all categories may have limited their chances of accurately stating the type that was praised. With a fewer number of categories, the effect of the vicarious praise may be stronger since a subject's attention is more likely to be directed to the praised category of response; and more subjects may be judged aware of the contingencies of reinforcement because of less complex modelling stimuli.

Representation of cognitive skill. The purpose of the present study was to show the effects of vicarious reinforcement of cognitive skill as exemplified by answering behaviours. This was achieved, but as discussed above, the results may have been more conclusive had a fewer number of categories of answers been used. Additionally, research using other representations of cognitive skill would be necessary to add strength to the theory of vicarious reinforcement of cognitive skill. The cognitive skill strengthened in the present study is limited to that exemplified by answering behaviours.

Frequency data. The data collected in the present investigation were of a frequency nature, limiting the statistical tests which could be used to detect the experimental effect to nonparametric ones. Stronger statistical tests can be used if data are ordinal, with a ranking between categories; or interval, with equal scale distance between values which represent the manifestations of the dependent variables.

As suggested above, if different representation of cognitive skill are used or different classification schemes are used for answering behaviours, it would be advantageous if the data were ordinal or interval in nature. That is, the manifestations of cognitive skill would ideally be on a scale from zero to the total number of skills, with equal distance between each of the skills. At the very least, the manifestations of cognitive skill could be ranked from lowest to highest. Had the Taxonomy been a true hierarchy, it would have gone further towards this condition than the Coding Schedule. With ordinal data, matching a specific cognitive skill would be most desirable but moving toward the cognitive skill would also be demonstrating a treatment effect. With interval data, not only the relative direction of movement, but the distance along the scale could be analyzed. More powerful parametric statistical procedures could then be used to determine whether a treatment effect existed.

Implications for Further Research

It can be seen from the preceding discussion of results and limitations that certain alterations of the present investigation would add to the volume of research dealing with vicarious reinforcement. The following suggestions for furthering the exploration of the vicarious reinforcement of cognitive skill are now summarized:

1. The use of more natural classroom settings to show the operation of vicarious reinforcers in classrooms is desirable. This, of course, would not allow for close control of the model and subject characteristics, but would illustrate the operation of vicarious reinforcers in natural settings with many competing stimuli.

2. Investigations of specific characteristics of the model and observers are desirable in order to explore more closely the effects when using models and observers of both sexes and to explore the effects of vicarious praise on subjects of different ages. Additional measures of observers' characteristics such as dependency and extroversion would add to the knowledge of types of observers that are likely to be influenced by vicarious praise.

3. Investigations of specific characteristics of the modelling stimuli with sufficiently large samples to allow for a more powerful tests are desirable. In this manner the influence of more intense praise and different positions of the praise could be investigated in more detail. Additionally, different types of potential reinforcers could be investigated.

4. Other representations of cognitive skill, including fewer categories of answering behaviours, could be formulated to demonstrate the generalized theory of vicarious reinforcement. These representations of cognitive skill could be arranged so they supply interval data in order that more powerful statistical tests could be used.

Summary of the Experiment

Because of the developmental nature of the study, a final summary may help to provide a total picture of the experimental portion of the present investigation. The study was concerned with effects of direct and vicarious reinforcers on cognitive skill operationalized as specific answering behaviours. Bandura's (1969) description of reinforcement provided the main theoretical background.

Cognitive skill, exemplified by answering behaviours, was dimensionalized by a modification of the Taxonomy of Educational Objectives (Bloom, 1956). The modified classification scheme, referred to as the Coding Schedule, had six categories of cognitive skill as summarized below:

1. Factualization - memorization
2. Comprehension - inference
3. Understanding - inference plus support
4. Conjecture - speculation
5. Application/Usage - statements of effect
6. Evaluation - judgmental statements

It was hypothesized that the praised category of answer would be strengthened for both the participant (by direct reinforcement) and the observer (by vicarious reinforcement).

Because of the low operant level of some categories of response from participants, and to allow for the manipulation of the specific category of answer which was provided and praised, a model was used. Additionally, the model was videotaped to allow for the exact replication of the modelling stimuli for each subject. The use of the model did not, however, allow for the exploration of direct reinforcers since the model followed a script giving predetermined responses.

The model and 53 subjects were fifth-form girls, aged fifteen to sixteen years. After watching a videotape of a model giving answers corresponding to all six categories of the Coding Schedule, with no praise given for any category of answer, a baseline measure of subjects' answering behaviour prior to the experimental treatment was taken. In the treatments which followed, each subject watched a model giving answers which corresponded to each of the six categories of the Coding Schedule. In these Treatment Videotapes, the model was praised for one specific category of answer. The model's answers were all in response to open-ended questions and were based on the content of a film, Reflections on Time (Encyclopaedia Britannica Educational Corporation, 1970).

In order to detect a treatment effect, the baseline measure (implying some preliminary measure of cognitive skill) was compared to the measure taken after a subject watched the selective praise of the model's answers on the Treatment Videotape. These measures taken after the Treatment Videotape were derived from open-ended questions. Subjects were asked one question to determine the type of answer they gave to a question similar to that asked of the model on the videotapes, (referred to as the Personal Question), one question to determine the type of answer a subject gave when answering for the model (referred to as the Empathetic Question), and other questions to determine subjects' awareness of the contingencies of reinforcement.

The data were analyzed according to the exact probability of subjects matching the category of answer that was praised on the Treatment Videotape and according to chi square analyses where the exact probabilities could not be calculated.

In summary the results showed:

1. Subjects gave answers of some categories more often than others when no praise was given to the model.
2. The order of presentation of the model's answers did not affect those of the subject.
3. Subjects matched the praised category of response more often than expected by chance alone.
4. Changes in intensity and position of praise did not produce greater treatment effects.
5. Subjects both aware and unaware of the contingencies of reinforcement were equally likely to match the praised category of response.

APPENDIX A

Short-Answer, Multiple-Choice and Oral Items Used
in Developmental Work

| | Page |
|---|------|
| 1. <u>Short-Answer Completion Items Arranged According to Their Cognitive Level of the Taxonomy of Educational Objectives</u> | 150 |
| 2. <u>Short-Answer Items as Originally Assigned to Subjects</u> | 152 |
| 3. <u>Revised Short-Answer Test (Form 1 to 6)</u> | 155 |
| 4. <u>Multiple-Choice Test (Form 1 and 2)</u> | 167 |
| 5. <u>Oral Test - Short Forms 1 and 2</u> | 173 |
| 6. <u>Oral Test - Long Forms 1 and 2</u> | 174 |

Short-Answer Completion Items Arranged
According to Their Cognitive Level of
the Taxonomy of Educational Objectives

| Level | Item |
|-------------------|--|
| 1. Knowledge | <p>Does the film define time? What are some of the rhythms and patterns of time that primitive man recognized? What was the film about? Name a clock of nature. How is time explored on geological explorations? What evidence of time is supplied by nature?</p> |
| 2. Comprehension. | <p>Explain what is meant by "A tenth of a second is a precious commodity." What does this cross-section of a tree say about the year corresponding to the indicated ring? Explain what is meant by "The past and present existing at the same time." Give examples of objective time as indicated by the film. Give examples of geological time as indicated by the film. Give examples of subjective time as indicated by the film.</p> |
| 3. Application | <p>If you were exploring a new land, how could you determine the age of the land? Show how a day may seem long to one person but short to another. Apply the principles in the film to define time. If there were no clocks around, how might you determine the approximate time? How could you use a tree to determine something about the year 1950? If you wanted to know the exact time, would you consult a wristwatch, cesium clock or sundial?</p> |

Short-Answer Completion Items cont.

| Level | Item |
|---------------|---|
| 4. Analysis | <p>Identify scenes in the film that illustrate the importance of time in modern man's life.</p> <p>Identify scenes in the film that illustrate the relativity of time.</p> <p>Identify scenes in the film that illustrate the constancy of time.</p> <p>Identify scenes in the film that illustrate man's awareness of time.</p> <p>Breakdown the content of the film into its main ideas.</p> <p>Identify scenes and script of the film that indicate the importance of time to earlier primitive man.</p> |
| 5. Synthesis | <p>Propose a statement about the end of time.</p> <p>Propose a statement about why you think time is an important phenomenon.</p> <p>How could you plan an experiment to show the relativity of time?</p> <p>Propose a statement about the importance of time in the life of an African tribesman.</p> <p>How could you plan an experiment to show the importance of time in modern man's life, briefly?</p> <p>How could you plan an experiment to show the inconsistent quality of time?</p> |
| 6. Evaluation | <p>Evaluate the use of the flashback in the film.</p> <p>Evaluate the effectiveness of the narration of the film.</p> <p>Evaluate the adequacy of geological explorations for determining time.</p> <p>Apply your own values to evaluate the film.</p> |

Short-Answer Items as Originally
Assigned to Subjects


Subject 1.

1. Which is more accurate, a cesium clock or sundial?
2. Paraphrase the expression "The clock of time is wound but once."
3. Draw a graph to show man's place in the world of time.
4. Identify scenes in the film that illustrate the unknown qualities of time.
5. Define time in your own words.
6. Evaluate the statement that "Time is as constant or variable as one makes it."

Subject 2.

1. Does the film define time?
2. Explain what is meant by "A tenth of a second is a precious commodity."
3. If you were exploring a new land, how could you determine the age of the land?
4. Identify scenes in the film that illustrate the importance of time in modern man's life.
5. Propose a statement about the end of time.
6. Evaluate the use of the flashback in the film.

Subject 3.

1. Does the film define time?
2. Evaluate the statement "Time is as constant or variable as one makes it."
3. What does this cross-section of a tree say about the year corresponding to the indicated ring? 
4. Propose a statement about the end of time.
5. If you were exploring a new land, how could you determine the age of the land?
6. Identify scenes in the film that illustrate the relativity of time.

Subject 4.

1. Explain what is meant by "The past and present existing at the same time."
2. Evaluate the effectiveness of the narration of the film.
3. Propose a statement about why you think time is an important phenomenon.
4. Show how a day may seem long to one person but short to another.
5. Give examples of objective time as indicated by the film.
6. How could you plan an experiment to show the relativity of time.
7. Identify scenes in the film that illustrate the constancy of time.
8. Identify scenes in the film that illustrate man's awareness of time.
9. Apply the principles in the film to define time.
10. Which is more accurate, a cesium clock or sundial?
11. What are some of the rhythms and patterns of time that primitive man recognized?
12. Paraphrase "The clock of time is wound but once."

Short-Answer Items as Originally Assigned cont.

Subject 5.

1. Explain what is meant by "A tenth of a second is a precious commodity."
2. Identify scenes in the film that illustrate the unknown qualities of time.
3. Evaluate the use of the flashbacks in the film.
4. How could you plan an experiment to show the importance of time in modern man's life.
5. If there were no clocks around, how might you determine the approximate time?
6. What was the film about?
7. Break down the content of the film into its main ideas.
8. How could you use a tree to determine something about the year 1950?
9. Name a clock of nature.
10. Propose a statement about the importance of time in the life of an African tribesman.
11. Give examples of geological time as indicated by the film.
12. Apply your own values to evaluate the film.
13. Define time in your own words.
14. Give examples of subjective time as indicated by the film.
15. How is time explored on geological explorations?
16. Evaluate the adequacy of geological explorations for determining time.
17. Draw a graph to show man's place in the world of time.
18. Identify scenes in the film that illustrate the importance of time in modern man's life.

Subject 6.


1. What evidence of time is supplied by nature?
2. Explain what is meant by "A tenth of a second is a precious commodity."
3. If there were no mechanical clocks around, how might you determine the approximate time?
4. Identify a few of the scenes in the film that illustrate the unknown qualities of time.
5. Propose a statement about the importance of time in the life of an African tribesman or other primitive man.
6. Apply your own values to evaluate the following statement:
Time is as constant or variable as one makes it.

Short-Answer Items as Originally Assigned cont.

Subject 7.

1. Name a clock of nature.
2. What was the film about, briefly?
3. What are some of the rhythms and patterns of time the primitive man recognized?
4. Give examples of objective time as indicated by the film.
5. If you wanted to know the exact time, would you consult a wristwatch, cesium clock or sundial?
6. Identify a few of the scenes in the film that illustrate man's awareness of time.
7. Propose an original statement about the possibility of the end of time.
8. Briefly, define time in your own words.
9. Apply your own values to evaluate the adequacy of geological explorations for determining time.

Subject 8.

1. Which is more accurate, a cesium clock or sundial?
2. How is time explored on geological explorations?
3. Does the film define time?
4. Give examples of geological time as indicated by the film.
5. Explain what is meant by "The past and present existing at the same time."
6. What does this cross-section of a tree say about the year corresponding to the indicated ring?
7. If you were exploring a new land, how could you determine the age of the land? 
8. Draw a graph to show man's place in the world of time.
9. How could you use a tree to determine something about the year 1950?
10. Identify scenes in the film that illustrate the constancy of time.
11. Breakdown the content of the film into its main ideas.
12. Identify scenes and script of the film that indicate the importance of time to earlier primitive man.
13. How could you plan an experiment to show the inconsistent quality of time?
14. Briefly, propose a statement about why you think time is an important phenomenon.
15. How could you plan an experiment to show the importance of time in modern man's life, briefly?
16. Apply your own values to evaluate the film.
17. Apply your own values to evaluate the use of the black and white flashback scenes in the film.
18. Apply your own values to evaluate the effectiveness of the narration of the film.

Revised Short-Answer Test (Form 1)

1. Propose a statement about the importance of time in the life of primitive man.
 2. Briefly, plan an experiment to show the importance of time in modern man's life.
 3. Give an example of the relativity of time.
 4. Briefly, propose a statement about the constancy of time.
-
1. Geology is to eons as _____ is to hours.
 2. What does this cross-section of a tree say about the year corresponding to the indicated ring?
 3. Give an example of how geological time might be used?
 4. What is meant by "Conquering time"?
-
1. Draw a graph to show man's place in the world of time.
 2. Primitive man was aware of time. How was his awareness different from ours?
 3. What quality of time is exemplified by the fact that 5 minutes is too long for a three-year-old to play a game, but not long enough for a thirteen-year-old?
 4. Drawing from the information given in the film, how will the Manawatu look in 2000 years?

Revised Short-Answer Test (Form 1) cont.

1. The film makes certain assumptions about time. What are a few of these assumptions?
 3. Identify scenes in the film that illustrate the constancy of time.
-
2. What was the aim or aims of the film?
 4. What is the relationship between objective and geological time?
-
1. What is Stonehenge?
 3. What was the film about?
-
2. Name a clock of nature.
 4. How are time and motion related?
-
1. Evaluate, and support your judgment, the appropriateness of the film for generations in 1000 years.
 3. Evaluate, and support your judgment, the value of time telling devices for cultural groups within New Zealand.
-
2. How relevant (and why) is the film to an African tribesman?
 4. Evaluate the statement: "I know what time it is." Support your answer.

Revised Short-Answer Test (Form 2)

1. Evaluate, and support your judgment, the appropriateness of the film for generations in 1000 years.
2. How relevant (and why) is the film to an African tribesman?
3. Evaluate, and support your judgment, the value of time telling devices for cultural groups within New Zealand.
4. Evaluate the statement: "I know what time it is." Support your answer.
1. Propose a statement about the importance of time in the life of primitive man.
3. Give an example of the relativity of time.
2. Briefly, plan an experiment to show the importance of time in modern man's life.
4. Briefly, propose a statement about the constancy of time.
1. What is Stonehenge?
3. What was the film about?
2. Name a clock of nature.
4. How are time and motion related?

Revised Short-Answer Test (Form 2) cont.

1. The film makes certain assumptions about time. What are a few of these assumptions?
 2. What was the aim or aims of the film?
 3. Identify scenes in the film that illustrate the constancy of time.
 4. What is the relationship between objective and geological time?
-
1. Geology is to eons as _____ is to hours.
 2. What does this cross-section of a tree say about the year corresponding to the indicated ring?
 3. Give an example of how geological time might be used.
 4. What is meant by "conquering time"?
-
1. Draw a graph to show man's place in the world of time.
 2. Primitive man was aware of time. How was his awareness different from ours?
 3. What quality of time is exemplified by the fact that **5 minutes** is too long for a three-year-old to play a game, but not long enough for a thirteen-year-old?
 4. Drawing from the information given in the film, how will the **Manawatu** look in 2000 years?

Revised Short-Answer Test (Form 3)

1. What is Stonehenge?
3. What was the film about?
2. Name a clock of nature.
4. How are time and motion related?
1. The film makes certain assumptions about time. What are a few of these assumptions.
3. Identify scenes in the film that illustrate the constancy of time.
2. What was the aim or aims of the film?
4. What is the relationship between objective and geological time?
1. Geology is to eons as _____ is to hours.
3. Give an example of how geological time might be used?
2. What does this cross-section of a tree say about the year corresponding to the indicated ring?
4. What is meant by "conquering time"?

Revised Short-Answer Test (Form 3) cont.

1. Propose a statement about the importance of time in the life of primitive man.
 2. Briefly, plan an experiment to show the importance of time in modern man's life.
 3. Give an example of the relativity of time.
 4. Briefly, propose a statement about the constancy of time.
-
1. Evaluate, and support your judgment, the appropriateness of the film for generations in 1000 years.
 2. How relevant (and why) is the film to an African tribesman?
 3. Evaluate, and support your judgment, the value of time telling devices for cultural groups within New Zealand.
 4. Evaluate the statement: "I know what time is." Support your answer.
-
1. Draw a graph to show man's place in the world of time.
 2. Primitive man was aware of time. How was his awareness different from ours?
 3. What quality of time is exemplified by the fact that 5 minutes is too long for a three-year-old to play a game, but not long enough for a thirteen-year-old?
 4. Drawing from the information given in the film, how will the Manawatu look in 2000 years?

Revised Short-Answer Test (Form 4)

1. Geology is to eons as _____ is to hours.
3. Give an example of how geological time might be used.
2. What does this cross-section of a tree say about the year corresponding to the indicated ring?
4. What is meant by "conquering time"?
1. The film makes certain assumptions about time. What are a few of these assumptions?
3. Identify scenes in the film that illustrate the constancy of time.
2. What was the aim or aims of the film?
4. What is the relationship between objectives and geological time?
1. Propose a statement about the importance of time in the life of primitive man.
3. Give an example of the relativity of time.
2. Briefly, plan an experiment to show the importance of time in modern man's life.
4. Briefly, propose a statement about the constancy of time.

Revised Short-Answer Test (Form 4) cont.

1. What is Stonehenge?
3. What was the film about?
2. Name a clock of nature.
4. How are time and motion related?
1. Draw a graph to show man's place in the world of time.
3. What quality of time is exemplified by the fact that 5 minutes is too long for a three-year-old to play a game, but not long enough for a thirteen-year-old?
2. Primitive man was aware of time. How was his awareness different from ours?
4. Drawing from the information given in the film, how will the Manawatu look in 2000 years?
1. Evaluate, and support your judgment, the appropriateness of the film for generations in 1000 years.
3. Evaluate, and support your judgment, the value of time telling devices for cultural groups within New Zealand.
2. How relevant (and why) is the film to an African tribesman?
4. Evaluate the statement: "I know what time it is." Support your answer.

Revised Short-Answer Test (Form 5)

1. Draw a graph to show man's place in the world of time.
 2. Primitive man was aware of time. How was his awareness different from ours?
 3. What quality of time is exemplified by the fact that 5 minutes is too long for a three-year-old to play a game, but not long enough for a thirteen-year-old?
 4. Drawing from the information given in the film, how will the Mānawatu look in 2000 years?
-
1. The film makes certain assumptions about time. What are a few of these assumptions?
 3. Identify scenes in the film that illustrate the constancy of time.
-
2. What was the aim or aims of the film?
 4. What is the relationship between objective and geological time?
-
1. Evaluate, and support your judgment, the appropriateness of the film for generations in 1000 years.
 3. Evaluate, and support your judgment, the value of time telling devices for cultural groups within New Zealand.
-
2. How relevant (and why) is the film to an African tribesman?
 4. Evaluate the statement: "I know what time it is." Support your answer.

Revised Short-Answer Test (Form 5) cont.

1. Propose a statement about the importance of time in the life of primitive man.
3. Give an example of the relativity of time.
2. Briefly, plan an experiment to show the importance of time in modern man's life.
4. Briefly, propose a statement about the constancy of time.
1. What is Stonehenge?
3. What was the film about?
2. Name a clock of nature.
4. How are time and motion related?
1. Geology is to eons as _____ is to hours.
3. Give an example of how geological time might be used?
2. What does this cross-section of a tree say about the year corresponding to the indicated ring?
4. What is meant by "conquering time"?

Revised Short-Answer Test (Form 6)

1. The film makes certain assumptions about time. What are a few of these assumptions?
3. Identify scenes in the film that illustrate the constancy of time.
2. What was the aim or aims of the film?
4. What is the relationship between objective and geological time?
1. Evaluate, and support your judgment, the appropriateness of the film for generations in 1000 years.
3. Evaluate, and support your judgment, the value of time telling devices for cultural groups within New Zealand.
2. How relevant (and why) is the film to an African tribesman?
4. Evaluate the statement: "I know what time it is." Support your answer.
1. Geology is to eons as _____ is to hours.
3. Give an example of how geological time might be used?
2. What does this cross-section of a tree say about the year corresponding to the indicated ring?
4. What is meant by "conquering time"?

Revised Short-Answer Test (Form 6) cont.

1. Propose a statement about the importance of time in the life of primitive man.
2. Briefly, plan an experiment to show the importance of time in modern man's life.
3. Give an example of the relativity of time.
4. Briefly, propose a statement about the constancy of time.
1. Draw a graph to show man's place in the world of time.
2. Primitive man was aware of time. How was his awareness different from ours?
3. What quality of time is exemplified by the fact that 5 minutes is too long for a three-year-old to play a game, but not long enough for a thirteen-year-old?
4. Drawing from the information given in the film, how will the Manawatu look in 2000 years?
1. What is Stonehenge?
2. Name a clock of nature.
3. What was the film about?
4. How are time and motion related?

Multiple-Choice Test (Form 1)

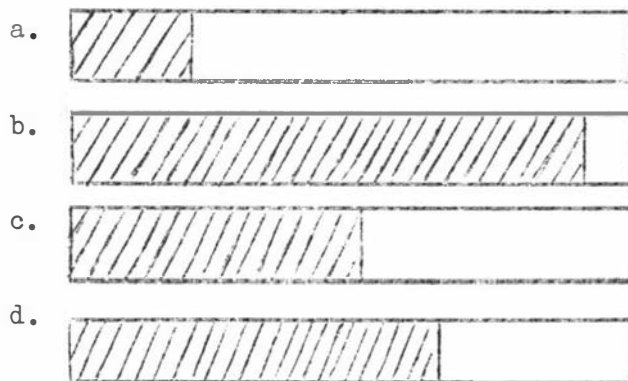
CIRCLE THE LETTER IN FRONT OF THE BEST ANSWER.

1. Which of the following best describes Stonehenge?
 - a. An early calendar
 - b. An early clock
 - c. An ancient sundial
 - d. An ancient superstition
2. Which statement of relationship is true?
 - a. Time and movement are the same
 - b. Time and movement are unrelated
 - c. Time and movement oppose each other
 - d. Time and movement define each other
3. Which of the following best describes how time is explored geologically?
 - a. Sand is poured from one container to another
 - b. Tree trunks are cut crosswise and the age is determined by the number of rings seen in the section
 - c. Pendulums determine the exact time of the earth's rotation
 - d. Rocks and fossils are used to determine the age of formations
4. Which term does not belong?
 - a. Pocket sundial
 - b. Cesium clock
 - c. Hourglass
 - d. Tree rings
5. Geology is to eons as _____ is to hours.
 - a. Minutes
 - b. Primitive man
 - c. Rhythms
 - d. Clock
6. Which of the following best describes what was meant by, "The past and present existing at the same time."
 - a. The old light of Adromeda illuminates our earth today
 - b. Remains of the past are still evident today
 - c. Grandparents and grandchildren exist at the same time
 - d. Old rocks determine the age of the present
7. In view of the ideas presented, which of the following would hold true?
 - a. Primitive man did not need objective time periods
 - b. Primitive man was unaware of the passage of time
 - c. Primitive man could predict seasons with the aid of Stonehenge
 - d. Primitive man was aware of only short time spans, like day and night

Multiple-Choice Test (Form 1) cont.

8. If a youngster approached and asked, "What is time?", the writer of the film would say:
- It is life
 - It is an unimportant phenomenon
 - It is a river that flows continuously
 - It is something that is difficult to define

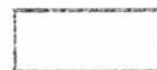
9. Which of the following graphs most accurately show man's place in the world of time?



KEY:



Time before the arrival of man



Time after the arrival of man

10. Which of the following is not an assumption made by the film?
- Time is constant
 - Time is variable
 - Time is defined
 - Time is important
11. Which of the following best describes the film?
- Objective and geological time are examined
 - Flashbacks are used to explore time
 - The Grand Canyon is a key to the past
 - Uses of Stonehenge are examined
12. Which of the following scenes illustrate the constancy of time?
- Scenes of ancient relics and remains
 - Scenes of tree rings being analyzed
 - Scenes of clocks, sundials and hourglasses
 - Scenes of primitive man

SUPPLY SHORT ANSWERS FOR THE FOLLOWING QUESTIONS.

13. How could you plan an experiment to show the importance of time in modern man's life, briefly.

Multiple-Choice Test (Form 1) cont.

14. Briefly, propose a statement about why you think time is an important phenomenon.

15. Propose an original statement about the possibility of the end of time.

16. Apply your own values and evaluate the film.

17. Apply your own values to evaluate the use of the black and white flashback scenes in the film.

18. Apply your own values to evaluate the effectiveness of the narration of the film.

Multiple-Choice Test (Form 2)

CIRCLE THE LETTER IN FRONT OF THE BEST ANSWER.

1. Which is more accurate?
 - a. A sundial
 - b. A cesium clock
 - c. An hourglass
 - d. A self-winding watch

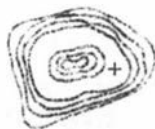
2. Which of the following best describes how time is explored geologically?
 - a. Sand is poured from one container to another
 - b. Tree trunks are cut crosswise and the age is determined by the number of rings seen in the section
 - c. Pendulums determine the exact time of the earth's rotation
 - d. Rocks and fossils are used to determine the age of the formation.

3. Which is the definition of time given by the film?
 - a. Time is an interval between two events
 - b. Time is not defined
 - c. Time is a phenomenon which governs our lives
 - d. Time is a fluid which flows continuously and constantly

4. Which of the following is not an example of objective time?
 - a. Sun time
 - b. Clock time
 - c. Cycles of the moon
 - d. Lifetime

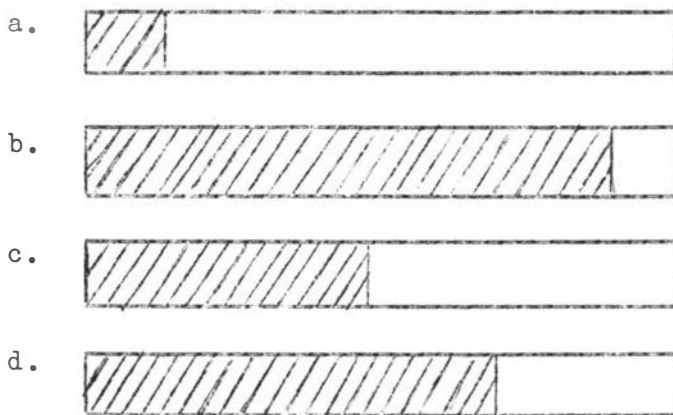
5. Which of the following best describes what was meant by : "The past and present existing at the same time."?
 - a. The old light of Adromeda illuminates our earth today
 - b. Remains of the past are still evident today
 - c. Grandparents and grandchildren exist at the same time
 - d. Old rocks determine the age of the present.

6. What does this cross-section of a tree say about the year corresponding to the indicated ring?
 - a. It was a cold year
 - b. It was a hot year
 - c. It was a wet year
 - d. It was a dry year



Multiple-Choice Test (Form 2) cont.

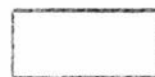
8. Which of the following graphs most accurately shows man's place in the world of time?



KEY:



Time before the arrival of man



Time after the arrival of man

9. If there were no mechanical clocks around, how might you best determine the approximate hour?
- By examining a tree trunk
 - By geological explorations
 - By looking at Stonehenge
 - By looking at the sun
10. Which of the following scenes illustrate the constancy of time?
- Scenes of ancient relics and remains
 - Scenes of tree rings being analyzed
 - Scenes of clocks, sundials and hourglasses
 - Scenes of primitive man
11. Which of the following is not one of the main ideas presented in the film?
- Geological time is used as a clue to the age of the universe
 - Stonehenge was used to tell the time of day
 - Man attempts to order and measure time
 - Time is both constant and variable
12. Which of the following scenes and script indicate the importance of time to early man?
- A cross-section of a tree
 - The geological exploration
 - Stonehenge
 - The swinging pendulum

SUPPLY SHORT ANSWERS FOR THE FOLLOWING QUESTIONS

13. How could you plan an experiment to show the importance of time in modern man's life, briefly.

Oral Test - Short Form 1

1. What evidence of time is supplied by nature?
2. Explain what the film meant by "A tenth of a second is a precious commodity."
3. How could you use a tree to determine something about the year 1950?
4. Identify a scene or bit of script that illustrates the unknown qualities of time.
5. Propose a statement about the importance of time in the life of primitive man.
6. Apply your own values to evaluate the following statement:
It is impossible to conquer time.

Oral Test - Short Form 2

1. What are some of the rhythms and patterns of time that primitive man recognized?
2. Give examples of geological time as indicated by the film.
3. If you wanted to know the exact time, what type of clock would you consult?
4. Identify a few of the scenes in the film that illustrate man's awareness of time.
5. Briefly, define time in your own words.
6. Apply your own values to evaluate the adequacy of geological explorations for determining time.

Oral Test - Long Form 1

1. What evidence of time is supplied by nature?
2. What was the most accurate clock mentioned in the film?
3. Explain what the film meant by "A tenth of a second is a precious commodity."
4. Give an example of objective time.
5. How could you use a tree to determine something about the year 1950?
6. If you were exploring a new land, how could you best determine the age of the land?
7. Identify a scene or bit of script that illustrates the unknown qualities of time.
8. What were the main ideas presented in the film?
9. Plan an experiment to show the inconsistent quality of time.
10. Briefly, propose a statement about why you think time is an important phenomenon.
11. Apply your own values to evaluate the following statement:
It is impossible to conquer time.
12. Apply your own values to evaluate the use of the black and white flashback scenes in the film.

Oral Test - Long Form 2

1. What are some of the rhythms and patterns of time that primitive man recognized?
2. What is the definition of time given by the film?
3. Give an example of geological time.
4. What is meant by "The past and present existing at the same time."?
5. What would the writer of the film say if a child asked "What is time?"?
6. If there were no mechanical clocks around, how might you determine the approximate hour?
7. Identify a few scenes in the film that illustrate man's awareness of time.
8. Identify a scene that illustrates the importance of time to primitive man.
9. Briefly, define time in your own words.
10. Propose an original statement about the possibility of the end of time.
11. Apply your own values to evaluate the adequacy of geological explorations for determining time.
12. Apply your own values to evaluate the effectiveness of the narration of the film.

APPENDIX B

Transcript of Open-Ended Questioning
of One Subject

1. Transcript

Page
176

Transcript

I = Investigator
S = Subject

- I.1 Would you tell us about the film
- S. Do you want my feelings or what it was about
- I.1 I want you to tell me your feelings
- S. It just posed a question - what's about time - it's all it posed but it didn't tell anybody anything. It didn't change or do anything or alter any knowledge anybody has. It just posed a question
- I.1 What did they have to consider from time to time. Obviously it's going to be a trap to give a leading question. It's your choice now - to see where we're going if these problems ... let's talk.... You said that it didn't really tell you anything - did you know it all?
- S. It didn't tell anything - except it did tell me - I didn't realize that moistures made rings on a tree widen - no, it didn't really tell me anything about the main thing it was about, time.
- I.1 Nevertheless, were you able to make up your mind at all - were you able to draw any conclusions about it
- S. Conclusions about what?
- I.1. Well, you said the main theme was ... the main question was the time
- S. mmmmm
- I.1 O.K., were you able to draw any conclusions?
- S. Well, I don't think there is an answer to that
- I.1 Do you think the film was trying to suggest there was an answer
- S. No, I think it was that it just tried to make people think - to see if they can come up with an answer - if they can come up with an answer they know what the meaning is
- I.1 If it was trying to make people think how did it do it
- S. Well, all the kind of stuff it - you know - it was all about - could mean about with time all about clocks and sundials and all that and it just kind of - at the end that was the most effective because it was all things old-fashioned - stone things and that and the sun and everything and there was cloud movements. And at the very beginning it just said, what was time - you know, very mysteriously.
- I.1 Why was it putting together so many different bits of information

Transcript cont.

- S. Kind of - to the evolution of man's thoughts of time, I would think.
- I.1 Uh, if somebody asked you the question 'what is time', what would you say
- S. Well, that I don't think it is anything as we know it. It is nothing solid and it's not really a thought but some people I suppose, would classify it as thought, but I couldn't.
- I.1 But yet we use the word - we talk about time
- S. Yes
- I.1 So it presumably has some meaning
- S. Mm - it's - I suppose in the way people use it as their overall view of time, history to man - right before Christ, right up to now and future. It's just a way of saying everything - I don't know - all times - that's wrong isn't it - it's kind of dates of time, and it's saying all these dates - it sounds weird but if ... I don't know what I'll say because even though people do use the word time, I think it is just a word - it's kind of got a meaning but it's just a meaning that nobody can put into words - they just know what it is
- I.1 Tell me, what way did the film try to get at the meaning
- S. Not just time - it was trying to make people not think of time - as just what you know how your watch - what the date is ...
- I.1 How did it do that
- S. How did the film do that?
- I.1 Yes
- S. It showed all sorts of the things - natural things of time, man's ~~thoughts of time and everything~~ - just time
- I.1 For example
- S. The Grand Canyon, pictures of the Grand Canyon and trees and everything that's happened through time - bits of nature
- I.1 What'll we ask her
- I.2 I don't know that - I know quite what I was going ...
It's almost unavoidable - the question doesn't mean anything, does it.
- I.1 Now tell me about what's going on afterwards
- I.2 Did you understand the narration of it - was it clear enough - you tell what he was saying
- S. I had to really listen hard to it you know I understood it
- I.1 Did you understand all the ideas that were generated ... all the words that were used
- S. Yes
- I.1 What's millennium

Transcript cont.

I.2 What's Stonehenge

S Pardon - Stonehenge - I don't know either

I.1 You don't

S. No, I don't think so. I might know it by another word - I don't mind...

I.1 Not that word you wouldn't - I am surprised, that you should know.

Would you like to ask some of the factual, some of the questions you have.

I.2 Yeah, I could- terribly organised - I hope it's coming out on the tape all right. Of course now I've got to pick these out because they were based on both parts - O.K. Leslie here goes

Here are some question. What evidence of time is supplied by nature?

S. That rock formations, the way - all the stuff that is formed, the rock and things is laid down in because that just couldn't happen at the snap of your fingers, and the trees

I.1 Do you mean just nature

S. Yes

I.2 What was the most accurate clock mentioned in the film?

S. Oh that - I've forgotten the name of it - that one that never loses a second - it's some great computer like thing - I can't think

I.1 How did it measure time - do you know - I mean anything that measures time has to have some sort of a reference - do you know what I mean

S. Oh yes - you mean like the sundial needs the sun as a reference

I.1 Yes - what was the mechanism in this clock?

S. I know it..

I.1 O.K. never mind

I.2 What did the film mean by the statement "a tenth of a second is a precious commodity"

S. Because that tenth of a second will never be there again - you just can't - you can't re-live something that happened a tenth of a second ago

I.2 O.K. Do you remember ... the film clip that was showing whenever that came on

S. No, short memory - eh?

I.2 It was the race that happened

S. Oh, yes -

I.2 I watched it too many times

I.1 That could be an occupational hazard

I.2 Yes, because to me that was perfectly obvious

Transcript cont.

- Can you identify any scene or bit of script that illustrates the unknown quantities of time?
- S. Oh yes the very beginning - you know there was a kind of pink sunrise I don't know how that's an unknown quality - unknown qualities. Can you re-phrase that question a bit.
- I.2 No, your answer was quite adequate. I'm sure I can't think of any better just the sheer vastness
- S. Oh yes - and that one of overall motion or something - it had the blue sky and it looked as though it was and it was a speed-up film of clouds moving
- I.2 That's it - what were the main ideas presented in that film clip
- S. You know it's just a question, you know - what's time - kind of what man has done to time and how man thinks he has conquered time, you know, like watches and that kind of stuff - what nature's kind of time - what nature has done
- I.2 What did you think of the black and white sections - the flashbacks
- S. Oh yes - I thought that was good because I think that was Kennedy making the speech, wasn't it. One of the Kennedys I suppose it would have been fairly modern for that kind of film and - but it was a shame it wasn't right up to the minute - that's what I thought was unfair it should have been Nixon or something you know
- I.1 He was in there
- S. Yeah, I know, I saw his big nose somewhere
I wasn't sure if it was him or Johnson. I think ...
- I.2 Why did you think it was good then?
- S. I like the way that it kept talking and what was it saying?
He said that before he started speaking the film said something that I thought the whole way through - those flashbacks. I've forgotten it now. You know, and I was thinking about that during the flashbacks and seeing how right it was - you know - that point it showed those flashbacks
- I.1 But you don't remember what they were
- S. No - great memory
- I.1 Can you work it out from the flashbacks
- S. I remember what it was - can man ever conquer time - you know - or make time go backwards

Transcript cont.

- I.1 How did you come to that - where ...
- S. I remember
- I.1 You just remember of you started to work it out backwards
- S. Bit of both I think because I started remembering that one kind of poor man at the very end - he was in colour and that ...
- I.2 What did you think the writer of the film would say if somebody asked him 'what is time'
- S. I think he'd say something like ... Something unpredictable or something unconceivable or something
- I.1 Why
- S. I don't know - because that - it didn't say what time was - it didn't say what time was and I don't think anybody - and I still don't think anybody can because I don't think you can give it precise right-on definition even in his own plan, I don't
- I.2 Could you pick out some scenes in the film that illustrates man's awareness of time.
- S. Oh yes, those high stone things - I didn't realise this but he said something that man did it to make the season or something so that each day of each year or some day of each year the sun came up just in the niche of that - that's show the pre-historic, whatever that means
- I.2 That was Stonchenge
- S. Was it - you're fussy
- I.2 Could you define time in your own words
- S. Well, he kind of - did
- I.1 No, can you
- S. You did ask me to do that, no I can't
- I.2 What are some of the rhythms and patterns of time that primitive man recognized
- S. Seasons - the moon phases - the position of the sun through the sky something to do with the stars - you know - different positions
- I.1 Day and night, perhaps
- S. That's obvious - even I notice that
- I.2 I've gone through all the questions
- I.1 Forgetting that you have seen the picture - if somebody then said to you - what's time - how would you answer

Transcript cont.

- S. I'd probably think about it for a while - I don't know - I still think I'd come to the same conclusion that I couldn't define it or I don't know - maybe I would - could say - I could have said so last year something - that's history - it's present - it's past and the future.
- I.1 Do you think that would satisfactorily explain the behaviour of two people when one says 'what's the time' and the other one says...
- S. That's the time - it's not 'what's time', they are completely different questions
- I.1 Are they
- S. Yes
- I.1 What's the difference
- S. What's the time - is saying what is the time at this moment or what is the marked time that man has come to recognize - that's what he's
8 saying
- I.1 I see - what's the other one
- S. It's just what is time - what is time - throughout history - in the future - every time
- I.1 O.K. thanks - this is going to make interesting listening I think I'll tell you what we're after then - I'd like to discuss it with you. People think in different ways about different things, now what we're after is trying to get(telephone break)
- I.2 see it 3 or 4 times I guess, it's so long
- I.1 We're after trying to get the different ways in which people think - were you conscious when we were asking you questions of having to think in different ways at all.
- S. Oh yes - cause you've done nothing - it was fairly obvious that you just weren't prepared for - what I was answering so you thought that I would - you were thinking of your own answers really - what your answer was, I would think.
- I.1 That's a very reasonable analysis but it doesn't actually answer my question which was - were you conscious of thinking differing kinds of answers that you were giving me
- S. Me thinking different kinds
- I.1 Yes, were you conscious of the fact that the kinds of answers that you gave differed - because they did from time to time
- S. Not really, it sounds goofy but I don't think I did

Transcript cont.

- I.1 No you are not making a judgment about them but you are not aware of the fact that they were different kinds of answers for instance, sometimes you gave bits of information that you saw in the film
- S. Oh yes - I gave different bits of information that came out of the analysis, in others I gave my opinion
- I.1 O.K.
- I.2 Hang on - other people's opinion or kind of tried to give other people's opinion
- I.1 Did you have difficulty with the kind of questions we asked you
- S. One of the ones you asked me at the beginning - I think it was
- I.1 Is that when you were giving the stall about what do I mean by the question
- S. Yes and there was one question that I asked you to re-phrase that I didn't - you didn't re-phrase
- I.1 Why did you do that
- S. I don't know but I don't think I really understood or didn't understand the wording or I...
- I.1 Did it
- S. I was thinking differently than I thought the way that you were thinking
- I.1 Did it have anything to do with your thinking about what kind of answer we might like to get
- S. I really don't think so. I was trying to baffle you actually
- I.1 All right - let's stop - let's have a little conference
- I.1 You saw the film. What do you think you could do now that you have seen the film.
- S. what you mean
- I.1 Can you use it - you just had a 7 minute experience with a film - can you use it
- S. Maybe I can - you know - some of the facts about that clock that never loses time and the actual - kind of the world - I don't know what the word is - the sun and everything the years lose 3 seconds so many years and all that stuff about the stone age man. Maybe I can use that for something but I don't know - maybe I can think about it now.

Transcript cont.

- I.2 Is there some way now that if you wanted to find something out about the weather of a particular year or the climate - the weather and the climates of a particular year, like the humidity in that - year, would you be able to do that now.
- S. Not more easily than I could before I don't think. Did I miss something?
- I.2 No, you have already told me that - I just ...
- I.1 Here again, I'm very conscious of the fact that the question give it away. Can you use it - that's sort of saying - does it make sense to ask you a question about using. Now I think that your answer says - I can't use it - but it is also an answer - it says 'Yes I know what your question is' so now are we after - and I can see how one could easily get entranced by the truth value of the answer - you would be worried about whether it is right or wrong and it is really incidental to our purposes I think - don't you
- I.2 The rightness or the wrongness
- I.1 The other think that I was conscious of too - and that was that - how much are we testing recall. In part we are sort of using the whole thing as a projected device - aren't we - a 7 minute projected - projection picture - and the extent of which the answers depend on recall will, in fact, effect the answers - it hasn't been a problem with this though.
- I.2 But to answer any of the categories of questions is there - there's got to be a certain amount of recall
- I.1 I expect
- S. I couldn't be ... very much. You have to think before it clicks over - ask about the question ... I think about ... and then suddenly remember it because of just ... a certain picture that was fairly relevant.
- I.2 The application bothers me a bit because it implies that she'll be given some kind of generalization to go from so that it's implicit.. and that's one I'm not sure that it comes out or perhaps looking at the transcript later. Well, for instance, whenever we asked her what the writer of the film would say about 'what's time' she would be replying something that she knew that the film had said or had not said about time. Synthesis - I think that that one came out pretty well, because obviously that she was synthesising evaluation and

Transcript cont.

analysis but once again I couldn't pinpoint really where those things were but the analysis... Maybe it.....the terms.

I.2 Broken them down - I'm sure.

APPENDIX C

Alternate Definitions of the Six Categories
of the Taxonomy of Educational Objectives

| | Page |
|---|------|
| 1. <u>Taxonomy as First Used to Classify Original Form of Responses to Open-ended Questions</u> | 186 |
| 2. <u>Taxonomy as Used to Classify Standardized Responses to Open-ended Questions</u> | 187 |
| 3. <u>Taxonomy as Used by Twenty Untrained Judges to Classify Ten Responses</u> | 188 |

Taxonomy as First Used to Classify Original
Form of Responses to Open-ended Questions

KNOWLEDGE

- The subject recalls facts and definitions, with minor alterations.
- The subject recognized information.
- The subject reproduces, with little or no change, what was presented.
- The subject recalls specifics, generalizations, methods, processes, pattern, structure, or setting.

COMPREHENSION

- The subject changes information into a different symbolic form or language.
- The subject paraphrases material or states in a different form.
- The subject performs in such a way that it is clear he knows what is being communicated.
- The subject explains or summarized a communication using a reordering, rearrangement or new view of the material.

APPLICATION

- The subject uses abstractions in particular and concrete situations.
The abstractions might be general ideas, rules of procedure, generalized method, principles or theories.
- The subject uses a method, rule, or principle to solve a new problem.
- The subject solves a lifelike problem that requires the identification of the issue and the selection and use of appropriate generalizations.

ANALYSIS

- The subject breaks down a communication into its constituent elements.
The analysis is intended to clarify the communication, indicate how the communication is organized, the way in which it manages to convey its effects, and/or its basis and management.
- The subject identifies elements in a communication.
- The subject identifies unstated assumptions in a communication.
- The subject points out interrelationships among the ideas presented.
- The subject identifies structural parts of the whole.

SYNTHESIS

- The subject puts together elements and parts to form a whole. The pattern or structure of the whole was not clearly there before.
- The subject develops a communication where he attempts to convey ideas, feelings and/or experiences.
- The subject develops a plan of work, or proposes ways of testing hypotheses.
The subject develops a set of abstract relations in order to explain or classify particular data.
- The subject solves a problem or offers a communication that requires original, creative thinking.
- The subject must design and create an original product.

EVALUATION

- The subject offers judgments about the value of material and methods for given purposes. The judgments may be quantitative and/or qualitative.
- The subject judges the accuracy of a communication from such evidence as logical accuracy and consistency.
- The subject sets up appropriate standards or value and determines how closely they are met.
- The subject judges good or bad, right or wrong according to standards he designates.

Taxonomy as Used to Classify
Standardized Responses to Open-ended Questions

- KNOWLEDGE: Statement of elements or components.
- COMPREHENSION: Statement of elements or components +
Implied relations or meaning.
- ANALYSIS: Statement of elements or components +
Implied relations or meaning +
Progressive reduction.
- SYNTHESIS: Stated or implied elements or components +
Implied relations or meanings +
Progressive production.
- APPLICATION: KNOWLEDGE, COMPREHENSION, ANALYSIS OR SYNTHESIS + use.
- EVALUATION: KNOWLEDGE, COMPREHENSION, ANALYSIS OR SYNTHESIS
+ valuing.

Taxonomy as Used by Twenty Untrained
Judges to Classify Ten Responses

1. KNOWLEDGE - Knowledge, as defined here, involves the recall of specifics and universals, the recall of methods and processes, or the recall of a pattern, structure, or setting. For measurement purposes, the recall situation involves little more than bringing to mind the appropriate material. Although some alteration of the material may be required, this is a relatively minor part of the task. The knowledge statements emphasize most the psychological processes of remembering. The process of relating is also involved in that a knowledge test situation requires the organization and reorganization of a problem such that it will furnish the appropriate signals and cues for the information and knowledge the individual possesses. To use an analogy, if one thinks of the mind as a file, the problem in a knowledge test situation is that of finding in the problem or task the appropriate signals, cues and clues which will most effectively bring out whatever knowledge is filed or stored.
2. COMPREHENSION - This represents the lowest level of understanding. It refers to a type of understanding or apprehension such that the individual knows what is being communicated and can make use of the material or idea being communicated without necessarily relating it to other material or seeing its fullest implications. Included in this is the translation, interpretation and/or extrapolation of material.
3. APPLICATION - The use of abstractions in particular and concrete situations. The abstractions may be in the form of general ideas, rules of procedures, or generalized methods. The abstractions may also be technical principles, ideas, and theories which must be remembered and applied.
4. ANALYSIS - The breakdown of a communication into its constituent elements or parts such that the relative hierarchy of ideas is made clear and/or the relations between the ideas expressed are made explicit. Such analyses are intended to clarify the communication, to indicate how the communication is organized, and the way in which it managed to convey its effects, as well as its basis and arrangement.
5. SYNTHESIS - The putting together of elements and parts so as to form a whole, This involves the process of working with pieces, parts, elements, etc., and arranging and combining them in such a way as to constitute a pattern or structure not clearly there before. It can be the production of a plan, a proposed set of operations or the production of a unique communication.
6. EVALUATION - Judgments about the value of material and methods for given purposes. Quantitative and qualitative judgments about the extent to which material and methods satisfy criteria. Use of a standard of appraisal. The criteria may be those determined by the respondent or those given to him.

APPENDIX D

Original and Final Form of the Coding
Schedule and Example Response Coding
Used to Train Judges

| | Page |
|---|------|
| 1. <u>Coding Schedule as First Formulated</u> | 190 |
| 2. <u>Examples of Responses in Each Category of the Coding Schedule</u> | 191 |
| 3. <u>Final Form of the Coding Schedule</u> | 192 |
| 4. <u>First Eighteen Responses Classified by Trained Judges</u> | 194 |
| 5. <u>Twenty Responses Classified by Trained Judges</u> | 195 |

Coding Schedule as First Formulated

1. FACTUALIZATION - A Factualization response, as defined here, involves the memorization or recall of facts and processes. No inferences are made. Cue words include:

That, as in "I saw that the car was"

2. COMPREHENSION - A comprehension response goes beyond what is factually presented and involves inference or intent. "It was trying to give examples of. . . ." and "The main idea of the poem was to. . . ." are examples of intent statements.

Cue words include:

How, as in "It showed how poor people live."

Aim to, as in "The story aimed to change my feelings."

3. UNDERSTANDING - An Understanding response involves an inference plus a reason for the inference as support.

Cue words include:

Because, as in "They went inside because of the rain."

Therefore, as in "The tree looked old, therefore it was cut down."

In order to, as in "He went away in order to make them angry."

By, as in "By talking about the weather, I felt cold."

4. CONJECTURE - A Conjecture response, as defined here, involves a speculation.

Cue words include:

If, as in "If the earth were flat"

I think, as in "I think he meant"

Note: The if of an if-statement is not necessarily positioned at the beginning of a response and, an I think must be followed by a speculation or nonrecall statement to be classed here.

5. APPLICATION/USAGE - An Application/Usage response involves the use or effect of something. It can be stated or implied statements of use or effect.

Cue words include:

Use, as in "I can use this to"

Made, as in "It made me feel sick."

Do, as in "It did change the temperature."

"Implied" do, as in "It changed the temperature." and
"I understand the poem."

6. EVALUATION - This involves a statement of value. A statement containing any mention of value is classed here if it is not specifically recall. Manneristic statements of how things are done are also classed here.

Cue words include:

Enjoy, as in "I enjoyed that."

Important, as in "School is of great importance."

To need, as in "It needs more light."

Quickly, as in "It was done quickly." (This tells in what "manner.")

Examples of Responses in Each Category
of the Coding Schedule

FACTUALIZATION

- It was showing the Grand Canyon, trees and bits of nature that have happened through time.
- It was showing evidence of time through nature, like rock formations.
- It was about time.
- It was saying that we can't use the rotation of the earth to measure time any more.

COMPREHENSION

- It was assuming that time is measurable, man can't control it and man has been aware of it since primitive days.
- It was trying to show time as a mystery.

UNDERSTANDING

- It was illustrating the unknown qualities of time by showing the pink sunrise and the overall motion of clouds.
- It was illustrating primitive man's awareness of time by showing Stonehenge.

CONJECTURE

- It was about motion and time, I think that to have motion you have to have time, but you don't necessarily need motion for time.
- It was presumably a universal film, but I think an African tribesman would be puzzled by the whole thing.
- It was about how we are guided by time, I think life must have been more relaxed when we didn't have watches and clocks.

APPLICATION/USAGE

- It was helping me think about the end of time.

EVALUATION

- It was hard to understand.
- It was good and fast moving
- It was interesting and enjoyable to watch the black and white section.

Final Form of the Coding Schedule

1. FACTUALIZATION - A Factualization response, as defined here, involves the memorization or recall of facts and processes. No inferences are made. This type of response is often subsumed in other categories.

That, as in "I saw that the car was . . ."
"It showed that. . ."

2. COMPREHENSION - A Comprehension response goes beyond what is factually presented and involves inference or intent. "It was trying to show some of the . . ." and "The main idea of the poem was to. . ." are examples of intent statements.

Cue words include:

How, as in "It showed how poor people live."
Aim to, as in "The story aimed to change my feelings."

3. UNDERSTANDING - An Understanding response involves an inference plus a reason for the inference as support.

Cue words include:

Because, as in "They went inside because of the rain."
Therefore, as in "The tree looked old, therefore it was cut down."
In order to, as in "He went away in order to make them angry."
By, as in "By talking about the weather, I felt cold."

4. CONJECTURE - A Conjecture response, as defined here involves a speculation.

Cue words include:

If, as in "If the earth were flat. . ."
I think, as in "I think he meant. . ."

NOTE: The if of an if-statement is not necessarily positioned at the beginning of a response. Also, a Conjectural if-statement usually has an explanation that sounds like a Comprehension or Understanding statement but is just the continuation of the if-statement. Thus, when a statement is typed as Conjectural often Factualization or Comprehension or Understanding are subsumed in it. Finally, an I think, must be followed by speculation or nonrecall to be categorized here. An I think plus simple recall is Factualization.

5. APPLICATION/USAGE - An Application/Usage response involves the use or effect of something. It can be stated or implied statements of use or effect.

Cue words include:

Use, as in "I can use this to. . ."
Made, as in "It made me feel sick."
Do, as in "It did change the temperature."
"Implied" do, as in "It changed the temperature."
"I understand the poem."

Final Form cont.

5. continued

NOTE: Use is categorized Application/Usage only when the respondent is indicating use beyond what was specifically presented otherwise it would be typed as Factualization or Comprehension or Understanding.

6. EVALUATION - This involves a statement of value. A statement containing any mention of value is categorized here, if it is not specifically recall. Manneristic statements of how things are done are also categorized here.
 Cue words include:

Enjoy, as in "I enjoyed that."

Important, as in "School is of great importance."

To need, as in "It needs more light."

Quickly, as in "It was done quickly," (This tells in what manner.)

Additional Classification Rules

- A negative statement is typed in the same category as the comparable positive statement.
- A response can be typed in more than one category. Each phrase that expresses a complete thought can be typed separately.
- A phrase can sometimes be reworded to ascertain the real meaning. Parenthetical phrases are sometimes eliminated if unnecessary to the main idea.
- Application means something has been done, not an intent to do. Intent is comprehension.
- Comprehension includes paraphrases.

First Eighteen Responses Classified
by Trained Judges

- It was making me think about time as the fourth dimension.
- It was illustrating the constancy of time by presenting the steady ticking of clocks.
- It was confusing and I didn't like the disjointed sense of time it showed.
- It was saying that you can't have time without motion.
- It was presenting many questions in order to get us to be aware of time.
- It was showing time as continuous movement of everything.
- It was explaining time, but I think if there were no people, there would be no concept of time.
- It was making me consider how I might be influenced by time.
- It was telling me that moisture makes rings on a tree widen.
- It was trying to show a short history of the measurement of time.
- It was defining time, but I don't think time is anything as we know it.
- It was about time and if someone asked the writer of the film for a definition of time, he would say it is unpredictable or unconceivable.
- It was mentioning a clock that never loses a second.
- It was hard to listen to.
- It was good the way it kept talking with the black and white scenes.
- It was aiming to encourage people to think about time and what it is.
- It was putting many bits of information together in order to show the evolution of man's thoughts about time.
- It was starting to change my attitude about time.

Twenty Responses Classified
by Trained Judges

1. _____ It was suggesting that man is getting better at measuring time.
2. _____ It was about time, I think time is history, past and present.
3. _____ It was saying conquering time but meant to beat a sport record.
4. _____ It was about time and I think the earth is part of time.
5. _____ It was aiming to get me to consider how time fixes itself.
6. _____ It was discussing time, I think time is happening and moving forward.
7. _____ It was trying to make people think of time as not just what your watch says.
8. _____ It was about time, if time is real I think it is just so man knows how to pace his life.
9. _____ It was showing all sorts of natural things.
10. _____ It was putting together many bits of information.
11. _____ It was hard to listen to.
12. _____ It was using colour and cloud movements to show time.
13. _____ It was talking about a previous tenth of a second which really meant breaking a record.
14. _____ It was making me think about how to run an experiment to find out the influence of time.
15. _____ It was illustrating the unknown qualities of time by showing a pink sunrise.
16. _____ It was questioning me about time.
17. _____ It was presenting the idea that time is a question.
18. _____ It was presenting questions about time.
19. _____ It was saying that Stonehenge marked the seasons.
20. _____ It was showing primitive man, if he really knew what time was we would be much further ahead today.

APPENDIX E

"Answers" as Classified Using the Taxonomy and
the Coding Schedule

| | Page |
|---|------|
| 1. <u>"Answers" Classified According to the Original Definition of the Taxonomy</u> | 197 |
| 2. <u>"Answers" Classified Using Revised Definition of the Taxonomy</u> | 200 |
| 3. <u>"Answers" Classified Using Both the Coding Schedule and Taxonomy</u> | 202 |

"Answers" Classified According to the
Original Definition of the Taxonomy

Knowledge

- It mentioned a clock that never loses a second, some great computer like thing.
- The main ideas presented were the question "What is time?", what man has done to time and how man thinks he has conquered time with watches, and nature's record of time.
- It showed the bits that we're all aware about it and then asked us what it would be like to control it and go back in time and finally, at the end, made us wonder about time with mysterious types of words.
- It showed Stonehenge, an ancient ring of large blocks of stone set up in England. It acts as a calendar of seasons.
- It was about time and how man has tried to harness time for his own purpose. It was asking what time is.
- It showed evidence of time through nature, like rock formations and trees.
- It showed the Grand Canyon and trees and bits of nature that have happened through time.
- It said we don't know much about time and we can't control it but we are getting better at measuring it.
- It showed rhythms and patterns that primitive man recognized. The rhythms and patterns included seasons, moon phases, the position of the sun and the stars.
- I thought it was about time mostly. How time goes, how it moves along and how things progress in time.
- It just posed a question - What is time? That's all it posed but it didn't tell anybody anything. It didn't change or do anything or alter any knowledge anybody has. It just posed a question.
- It showed Stonehenge. That's a place where some of the older type of people, before Christ, set up a temple sort of thing to determine seasons.
- It was trying to make people think by showing clocks and sundials, old-fashioned stone things, the sun and cloud movements. Also, at the very beginning it asked "What is time?" very mysteriously.
- It was showing time as continuous movement of everything.
- It told me that as a tree grows, each year another layer is added to the trunk. You can tell how old a tree is by counting the layers.
- It started off showing that primitive man was aware of time and how he measured it, such as Stonehenge. The pictures made a lot of the sun and it showed how man divided days into hours into minutes into seconds and how man became more accurate in measuring time. Finally, towards the end, it suggested and sort of posed a question, "What would it be like if you could control time, and right at the end it asked what time was.

"Answers" Classified According to
Original Definition cont.

Comprehension

- It said that primitive man used time for planting crops in the correct seasons, for waking and sleeping, natural things. We use it for a 9 to 5 job, working so many hours on a particular job. Primitive man hasn't divided time up so much because he has more time for other things and has so much to do.
- The film assumed that time is what passes and isn't really anything particular. It is what man has made in order to determine what has passed.
- The main idea of the film was to tell people watching it that man is some sub-race. He can't harness nature for his own use.
- Based on what the film showed, if someone asked me what time was, I would say it is the world progressing and different things happening.
- It illustrated man's awareness of time by showing Stonehenge that was used to predict the season.
- It just tried to make people think, to see if they can come up with an answer. If they can come up with an answer, they know what the meaning is.
- It showed primitive man right at the beginning. I think he was aware of time by noticing the periods of darkness and sensing something was going by, something was left in the past.
- It illustrated the constancy of time by presenting the steady ticking of clocks.
- Man's awareness of time came through in the narration.
- The writer of the film would say time is unpredictable or unconceivable if someone asked for a definition.
- It talked about conquering time, which means being able to control it and move around in it, apart from thoughts that involve the flow of it.
- It wants people to think about time and become aware of the extent to which they are governed by it and the very limited knowledge that man has about it.
- It was putting many bits of information together to show the evolution of man's thoughts about time.
- It was trying to make people think of time as not just what your watch says or the date it is.
- I think it was trying to say that time is something we can't get a hold of.
- The aim of the film was to get people to think about time and about what it is.
- It illustrated the unknown qualities of time with the pink sunrise, and the overall motion of the clouds moving across the blue sky.
- It assumed that time is measurable, man can't control it yet man has been aware of it ever since primitive days.
- "A tenth of a second is a precious commodity" means that a tenth of a second will never be there again. You can't relive something that happened a tenth of a second ago.

"Answers" Classified According to
Original Definition cont.

Application

- If I was doing a school project on what time is I might be able to use some of the ideas in the film.

Analysis

- It used pictures to get points across. Like the pictures of past events and the clocks to show the measurement of time. It used fast motions for the clouds streaming across the sky. Time was actually there, therefore it takes a while to perceive the moment. You could just glance at the picture and tell time was moving across.

Synthesis

- It made me think about time as the fourth dimension.
- It said you can't have time without motion. I think that to have motion you have to have time, but you don't necessarily need motion for time.
- It was about time. Time is equally important to primitive man and modern man because both cannot exist without it.
- An African tribesman, presuming he knew the language of the film, would be puzzled about the people who made the film and would wonder what it is about and why they make such a big deal about time. Because tribesmen usually don't have clocks and they have more time to spend on a thing, they don't need to divide time up so much.

Evaluation

- The black and white sections were good. It was Kennedy making a speech and that indicated that it was a fairly modern kind of film.
- I liked the way it kept talking with the black and white scenes and illustrated the idea that man can make time go backwards.
- I enjoyed seeing how different clocks work.
- The narration of the film was poor because some of the words were unclear and some of the words were more for adults, people who are going to university.
- It changed scenes quickly. It was hard to get hold of what was happening.

Other

- The Stonehenge thing, I didn't understand that.
- I don't think time is anything as we know it. It is nothing solid and it's not really a thought. Some people would classify it as a thought, I suppose, but I couldn't.
- I had to really listen hard to it but I understood it.

"Answers" Classified Using Revised
Definition of the Taxonomy

Knowledge

- It was about time.
- It was showing Stonehenge, rocks, trees, a primitive man and a race.
- It was showing a lot of clocks and hourglasses and sundials.
- It was saying the cesium clock is very accurate, more accurate than mechanical ones.
- It was saying that a tenth of a second counts, especially in a race.
- It was saying that we can't use the rotation of the earth to measure time anymore because it loses 3 seconds in some years.

Comprehension

- It was talking about "conquering time" which means beating a sport record.
- It was saying "splitting seconds" but really meant breaking a second into smaller parts with more precise clocks.
- It was implying that layers of rock measure millions of years as clocks measure hours.
- It was indicating that primitive man was aware of time in bigger units (like days and seasons) than we are.
- It was showing a brief history of the measurement of time.
- It was telling how the units with which we measure time have got smaller and more accurate over the years.

Application

- It was showing me that I can use a tree trunk to find out how wet it was in some years.
- It was telling how clocks are inaccurate. Because of that, I know now that I wouldn't be able to tell someone the definite time.
- It was telling me that I could determine the age of a new land by examining layers of rock. Maybe I could determine the age of a new planet by this method too.
- It was concerned about primitive man's concept of time. Like the primitive man I saw, I could tell when to plan crops by observing nature if I was in a strange or remote place.
- It was presenting time in nature. I suppose those layers of rock go back to before there were written records. Scientists could use them to learn about past things like floods and ice ages.

"Answers" Classified Using Revised
Definition cont.

Analysis

- It was making assumptions about time. It assumed that time exists, it can be measured, it is continuous, and we are all affected by it.
- It was using the black and white scenes to give a different view of time, I think. It was using them to show something besides forward time.
- It was organized into 3 sections. First it showed different ways of measuring time to give a history to it. Next, it said we still don't know what time is and, finally, it explored the possibility of going backwards in time by using some clever camera work.
- It was showing a race to illustrate one of the reasons why man separates time into small units.
- It was presenting old time pieces, Stonehenge, and primitive man to show that man has tried to know and understand time for centuries.
- It was an inquiry film because it presented many questions in order to get the audience to think about time.

Synthesis

- It was concerned about peoples' awareness of time. I think life must have been more relaxed when they didn't have watches. Now, because of them, we have to be more accurate about meeting people.
- It was making me think of time as the fourth dimension and to wonder about ways of getting into that dimension and moving both backwards and forwards in it.
- It was making me think about how to run an experiment to tell how important time is to man. You could remove all the clocks and watch what happens.
- It was influencing me in such a way that, if I had to define time now, I would say it is an unknown occurrence that influences my life.
- It was helping me think about the end of time. If there were no people, there would be no concept of time.
- It was in need of a section about the unsureness of time and how time can seem both long and short.

Evaluation

- It was hard to understand because the man used words I didn't know. I didn't like that.
- It was hard to listen to and I didn't like the sound track. The sound track was bad because sometimes the music covered up the man's voice.
- It was confusing and I didn't like the film because it gave a disjointed sense of time. The pictures often didn't really go along with the words.
- It was good because I liked the way it held my interest by keeping the scenes constantly moving.
- It was good because the black and white flashbacks showed how strange it would be to turn back time.
- It was poor in that it doesn't apply to everyone. Some people, like hippies and primitive men, aren't concerned about time.

"Answers" Classified Using Both the
Coding Schedule and Taxonomy

Please read the CLASSIFICATION SCHEME carefully and classify each response into one category. Designate the category by 1, 2, 3, 4, 5, or 6 in the space provided. The question "Tell me about the film." prompted each response from a school child.

CATEGORY

_____ I had to listen very hard.

_____ It was showing time as continuous movement of everything.

_____ It didn't change or do anything to alter any knowledge anybody has.

_____ It mentioned a clock that never loses a second.

_____ It told me that moisture makes rings on a tree widen.

_____ It was putting many bits of information together in order to show the evolution of man's thoughts about time.

_____ It made me think about time as the fourth dimension.

_____ I don't think time is anything as we know it.

_____ The aim of the film was to get people to think about time and about what it is.

_____ I liked the way it kept talking with the black and white scenes.

_____ I think that to have motion you have to have time, but you don't necessarily need motion for time.

_____ It illustrated the constancy of time by presenting the steady ticking of clocks.

_____ It said that you can't have time without motion.

_____ The writer of the film would say time is unpredictable or unconceivable if someone asked for a definition.

_____ The narration of the film was poor.

APPENDIX F

Scripts for Videotapes and Experimental
Sessions

| | Page |
|---|------|
| 1. <u>Scripts for Baseline and Treatment Videotapes</u> | 204 |
| 2. <u>Investigator's Script for the Experiment</u> | 214 |

Scripts of Baseline and Treatment Videotapes

BASELINE A SCRIPT

- E - I'm going to ask you about the movie you have just seen, Fiona.
Think carefully and answer as best you can.
- E - Fiona, tell me about the film.
- F - It was saying that you can't have time without motion.
- E - O.K.
- E - Fiona, tell me about the film.
- F - It was making me consider how I might be affected by time.
- E - O.K.
- E - Fiona, tell me about the film.
- F - It was trying to show a short history of the measurement of time.
- E - O.K.
- E - Fiona, tell me about the film.
- F - It was illustrating the constancy of time by presenting the steady ticking of clocks.
- E - O.K.
- E - Fiona, tell me about the film.
- F - It was defining time, but I don't think time is anything as we know it.
- E - O.K.
- E - Fiona, tell me about the film.
- F - it was hard to listen to.
- E - O.K.

BASELINE B SCRIPT

- E - I'm going to ask you about the movie you have just seen, Fiona.
Think carefully and answer as best you can.
- E - Fiona, tell me about the film.
- F - It was trying to show a short history of the measurement of time.
- E - O.K.
- E - Fiona, tell me about the film.
- F - It was hard to listen to.
- E - O.K.
- E - Fiona, tell me about the film.
- F - It was illustrating the constancy of time by presenting the steady ticking of clocks.
- E - O.K.
- E - Fiona, tell me about the film.
- F - It was defining time, but I don't think time is anything as we know it.
- E - O.K.

Baseline and Treatment Scripts cont.

BASELINE B SCRIPT cont.

E - Fiona, tell me about the film.

F - It was saying that you can't have time without motion.

E - O.K.

E - Fiona, tell me about the film.

F - It was making me consider how I might be affected by time.

E - O.K.

BASELINE C SCRIPT

E - I'm going to ask you about the movie you have just seen, Fiona.
Think carefully and answer as best you can.

E - Fiona, tell me about the film.

F - It was illustrating the constancy of time by presenting the steady ticking of clocks.

E - O.K.

E - Fiona, tell me about the film.

F - It was saying that you can't have time without motion.

E - O.K.

E - Fiona, tell me about the film.

F - It was hard to listen to.

E - O.K.

E - Fiona, tell me about the film.

F - It was trying to show a short history of the measurement of time.

E - O.K.

E - Fiona, tell me about the film.

F - It was making me consider how I might be affected by time.

E - O.K.

E - Fiona, tell me about the film.

F - It was defining time, but I don't think time is anything as we know it.

E - O.K.

BASELINE D SCRIPT

E - I'm going to ask you about the movie you have just seen, Fiona.
Think carefully and answer as best you can.

E - Fiona, tell me about the film.

F - It was defining time, but I don't think time is anything as we know it.

E - O.K.

E - Fiona, tell me about the film.

F - It was illustrating the constancy of time by presenting the steady ticking of clocks.

E - O.K.

E - Fiona, tell me about the film.

F - It was making me consider how I might be affected by time.

E - O.K.

Baseline and Treatment Scripts cont.

BASELINE D SCRIPT cont.

E - Fiona, tell me about the film.

F - It was hard to listen to.

E - O.K.

E - Fiona, tell me about the film.

F - It was saying that you can't have time without motion.

E - O.K.

E - Fiona, tell me about the film.

F - It was trying to show a short history of the measurement of time.

E - O.K.

BASELINE E SCRIPT

E - I'm going to ask you about the movie you have just seen, Fiona.
Think carefully and answer as best you can.

E - Fiona, tell me about the film.

F - It was making me consider how I might be affected by time.

E - O.K.

E - Fiona, tell me about the film.

F - It was hard to listen to.

E - O.K.

E - Fiona, tell me about the film.

F - It was saying that you can't have time without motion.

E - O.K.

E - Fiona, tell me about the film.

F - It was trying to show a short history of the measurement of time.

E - O.K.

E - Fiona, tell me about the film.

F - It was defining time, but I don't think time is anything as we know it.

E - O.K.

E - Fiona, tell me about the film.

F - It was illustrating the constancy of time by presenting the steady ticking of clocks.

E - O.K.

Baseline Treatment Scripts cont.

BASELINE F SCRIPT

E - I'm going to ask you about the movie you have just seen, Fiona.
Think carefully and answer as best you can.

E - Fiona, tell me about the film.

F - It was hard to listen to

E - O.K.

E - Fiona, tell me about the film.

F - It was defining time, but I don't think time is anything as we know it.

E - O.K.

E - Fiona, tell me about the film.

F - It was illustrating the constancy of time by presenting the steady ticking of clocks.

E - O.K.

E - Fiona, tell me about the film.

F - It was making me consider how I might be affected by time.

E - O.K.

E - Fiona, tell me about the film.

F - It was trying to show a short history of the measurement of time.

E - O.K.

E - Fiona, tell me about the film.

F - It was saying that you can't have time without motion.

E - O.K.

Baseline and Treatment Scripts cont.

TREATMENT A SCRIPT

E - Now I'm going to ask you more questions about the movie. Once again you should think carefully and answer as best you can. Take as much time as you need. There is no reason to hurry.

E - Fiona, can you tell me some more about the film.

F - It was starting to change my attitude about time.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was confusing and I didn't like the disjointed sense of time it showed.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was telling me that moisture makes rings on a tree widen.

Reinforcement : E - That's a very good answer.

Reinforcement and reason: E - That's a very good answer. You have remembered what the film said.

E - Fiona, can you tell me some more about the film.

F - It was aiming to encourage people to think about time and what it is.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was presenting many questions in order to get us to be aware of time.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was about time, if time is real, I think it is just so man knows how to pace his life.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was making me think about how to run an experiment to find out about the influence of time.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was good the way it kept talking with the black and white scenes.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was mentioning a clock that never loses a second.

Reinforcement : E - That's a very good answer

Reinforcement and reason: E - That's a very good answer. You have remembered what the film said.

E - Fiona, can you tell me some more about the film.

F - It was showing time as continuous movement of everything.

E - O.K.

E - Fiona, can you tell me more about the film.

F - It was putting many bits of information together in order to show the evolution of man's thoughts about time.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was explaining time, but I think if there were no people there would be no concept of time.

E - O.K.

Baseline and Treatment Scripts cont.

TREATMENT B SCRIPT

E - Now I'm going to ask you more questions about the movie. Once again you should think carefully and answer as best you can. Take as much time as you need. There is no reason to hurry.

E - Fiona, can you tell me some more about the film.

F - It was confusing and I didn't like the disjointed sense of time it showed.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was telling me that moisture makes rings on a tree widen.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was aiming to encourage people to think about time and what it is.

Reinforcement: E - That's a very good answer.

Reinforcement and reason: E - That's a very good answer. You have made an inference from what the film showed.

E - Fiona, can you tell me some more about the film.

F - It was presenting many questions in order to get us to be aware of time.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was about time, if time is real, I think it is just so man knows how to pace his life.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was starting to change my attitude about time.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was good the way it kept talking with the black and white scenes.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was mentioning a clock that never loses a second.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was showing time as continuous movement of everything.

Reinforcement: E - That a very good answer.

Reinforcement and reason: E - That's a very good answer. You have made an inference from what the film showed.

E - Fiona, can you tell me some more about the film.

F - It was putting many bits of information together in order to show the evolution of man's thoughts about time.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was explaining time, but I think if there were no people there would be no concept of time.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was making me think about how to run an experiment to find out about the influence of time.

E - O.K.

Baseline and Treatment Scripts cont.

TREATMENT C SCRIPT

E - Now I'm going to ask you more questions about the movie. Once again you should think carefully and answer as best you can. Take as much time as you need. There is no reason to hurry

E - Fiona, can you tell me some more about the film.

F - It was telling me that moisture makes rings on a tree widen.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was aiming to encourage people to think about time and what it is.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was presenting many questions in order to get us to be aware of time.

Reinforcement: E - That's a very good answer.

Reinforcement and reason: E - That's a very good answer. You have made an inference from what the film said and you have given a reason for the inference.

E - Fiona, can you tell me some more about the film.

F - It was about time, if time is real, I think it is just so man knows how to pace his life.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was starting to change my attitude about time.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was confusing and I didn't like the disjointed sense of time it showed.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was mentioning a clock that never loses a second.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was showing time as continuous movement of everything.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was putting many bits of information together in order to show the evolution of man's thoughts about time.

Reinforcement: E - That's a very good answer.

Reinforcement and reason: E - That's a very good answer. You have made an inference from what the film showed and you have given a reason for the inference.

E - Fiona, can you tell me some more about the film.

F - It was explaining time, but I think if there were no people there would be no concept of time.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was making me think about how to run an experiment to find out about the influence of time.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was good the way it kept talking with the black and white scenes.

E - O.K.

Baseline and Treatment Scripts cont.

TREATMENT D SCRIPT

- E - Now I'm going to ask you more questions about the movie. Once again you should think carefully and answer as best you can. Take as much time as you need. There is no reason to hurry.
- E - Fiona, can you tell me some more about the film.
- F - It was aiming to encourage people to think about time and what it is.
- E - O.K.
- E - Fiona, can you tell me some more about the film.
- F - It was presenting many questions in order to get us to be aware of time.
- E - O.K.
- E - Fiona, can you tell me some more about the film.
- F - It was about time, if time is real, I think it is just so man knows how to pace his life.
- Reinforcement: E - That's a very good answer.
- Reinforcement and reason: E - That's a very good answer. You have speculated about what you saw.
- E - Fiona, can you tell me some more about the film.
- F - It was starting to change my attitude about time.
- E - O.K.
- E - Fiona, can you tell me some more about the film.
- F - It was confusing and I didn't like the disjointed scenes of time it showed.
- E - O.K.
- E - Fiona, can you tell me some more about the film.
- F - It was telling me that moisture makes rings on a tree widen.
- E - O.K.
- E - Fiona, can you tell me some more about the film.
- F - It was showing time as continuous movement of everything.
- E - O.K.
- E - Fiona, can you tell me some more about the film.
- F - It was putting many bits of information together in order to show the evolution of man's thoughts about time.
- E - O.K.
- E - Fiona, can you tell me some more about the film.
- F - It was explaining time, but I think if there were no people there would be no concept of time.
- Reinforcement: E - That's a very good answer.
- Reinforcement and reason: E - That's a very good answer. You have speculated about what you saw.
- E - Fiona, can you tell me some more about the film.
- F - It was making me think about how to run an experiment to find out about the influence of time.
- E - O.K.
- E - Fiona, can you tell me some more about the film.
- F - It was good the way it kept talking with the black and white scenes.
- E - O.K.
- E - Fiona, can you tell me some more about the film.
- F - It was mentioning a clock that never loses a second.
- E - O.K.

Baseline and Treatment Scripts cont.

TREATMENT E SCRIPT

E - Now I'm going to ask you more questions about the movie. Once again you should think carefully and answer as best you can. Take as much time as you need. There is no reason to hurry.

E - Fiona, can you tell me some more about the film.

F - It was presenting many questions in order to get us to be aware of time.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was about time, if time is real, I think it is just so man knows how to pace his life.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was starting to change my attitude about time.

Reinforcement: E - that's a very good answer.

Reinforcement and reason: E - That's a very good answer. You have applied what was presented.

E - Fiona, can you tell me some more about the film.

F - It was confusing and I didn't like the disjointed sense of time it showed.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was telling me that moisture makes rings on a tree widen.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was aiming to encourage people to think about time and what it is.

E - O.K.

E - Fiona, can you tell some more about the film.

F - It was putting many bits of information together in order to show the evolution of man's thoughts about time.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was explaining time, but I think if there were no people there would be no concept of time.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was making me think about how to run an experiment to find out about the influence of time.

Reinforcement: E - That's a very good answer.

Reinforcement and reason: E - That's a very good answer. You have applied what was presented.

E - Fiona, can you tell me some more about the film.

F - It was good the way it kept talking with the black and white scenes.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was showing time as continuous movement of everything

E - O.K.

Baseline and Treatment Scripts cont.

TREATMENT F SCRIPT

E - Now I'm going to ask you more questions about the movie. Once again you should think carefully and answer as best you can. Take as much time as you need. There is no reason to hurry.

E - Fiona, can you tell me some more about the film.

F - It was about time, if time is real, I think it is just so man knows how to pace his life.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was starting to change my attitude about time.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was confusing and I didn't like the disjointed sense of time it showed.

Reinforcement: E - That's a very good answer.

Reinforcement and reason: E - That's a very good answer. You have made a value judgment concerning the film.

E - Fiona, can you tell me some more about the film.

F - It was telling me that moisture makes rings on a tree widen.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was aiming to encourage people to think about time and what it is.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was presenting many questions in order to get us to be aware of time.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was explaining time, but I think if there were no people there would be no concept of time.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was making me think about how to run an experiment to find out about the influence of time.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was good the way it kept talking with the black and white scenes.

Reinforcement: E - That's a very good answer.

Reinforcement and reason: E - That's a very good answer. You have made a value judgment concerning the film.

E - Fiona, can you tell me some more about the film.

F - It was mentioning a clock that never loses a second

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was showing time as continuous movement of everything.

E - O.K.

E - Fiona, can you tell me some more about the film.

F - It was putting many bits of information together in order to show the evolution of man's thoughts about time.

E - O.K.

Investigator's Script for the Experiment

Experimenter - You are probably curious as to what will happen today, so let me tell you. I'm interested in how people react to both television and film. People spend many hours watching television and films but little is known about how they affect us. You can help me find out how people react to films and television by watching both and by answering some simple questions. First I'd like you to watch and listen carefully to a bit of film.

----- FILM 1 -----

Experimenter - Now we'll see a videotape of a girl named Fiona being asked the same question on several occasions. She is being questioned about the bit of film you just saw. Watch and listen carefully.

----- BASELINE VIDEOTAPE -----

Experimenter - Fiona gave six good answers to the question. Here is a list of all the answers she gave. From these six, I'd like you to choose the one that is nearest to the kind of answer you would have given to the question "Can you tell me about the film?". Put an X beside the one you choose and give it to me when you have finished.

(WAIT TO COLLECT PAPER)

Experimenter - Now we'll see Fiona being asked the same question on several occasions once again. I'd like you to watch and listen carefully.

----- TREATMENT VIDEOTAPE -----

Experimenter - We have heard Fiona give many answers about the film. Now I'd like you to watch and listen to another bit of film and you will have a chance to answer questions about it.

----- FILM 2 -----

Experimenter - I'll turn on the recorder now so that I have a record of your answers. There are just a few questions about this last piece of film for you to answer. First, what would you answer to the question "Can you tell me about the film?" Can you tell me why you answered that way?

The next question is, what do you think Fiona, the girl on the videotape would answer to the question "Can you tell me about the film?" Why do you think she would say that ??

Did you notice any difference in the types of answers that Fiona was giving?

Did you notice anything different in what I was saying to Fiona on the videotape. Why do you think I said that?

It's most important that you don't tell anybody, especially 5th Form girls what we did today. If someone does ask, simply tell them that you watched a movie and a videotape and answered some question. Of course, if your parents enquire, you can tell them what happened.

APPENDIX G

Transcript of Questioning of a Subject
after Experimental Treatment

| | Page |
|------------------------------|------|
| 1. Transcript from Subject 4 | 216 |

Transcript from Subject 4

- E - There are just a few questions about this last piece of film for you to answer. First, what would you answer to the question "Can you tell me about the film?"?
- S - It was telling about how time had to do with history, different parts of the rocks and how far it goes back. And where man has began, and different light sent to the earth, and it said only man began.
- E - O.K. Can you tell me why you'd answer that way? What made you think of those particular things?
- S - Those are the ones that stood out. It was trying to get across to us the history of the earth and other parts of the outer galaxies.
- E - O.K. The next question is: What do you think Fiona, the girl on the videotape would answer to the question "Can you tell me about the film?"?
- S - She'd say something about the rocks and the man and how he's knocking at the rocks. The colours of the film, I think.
- E - O.K. Why do you think she'd say those things.
- S - Those are the most striking things.
- E - Those are the most striking things. O.K. Good. Did you notice any different kinds of answers that Fiona was giving on either of the videotapes? Did you notice any different kinds of answers?
- S - Yes, the first one. She may have seen it but it mightn't have been clear as the second time. She might have thought about it a bit more in the second videotape.
- E - O.K. So you think that she gave different kind of answers or they were more clear the second time.
- S - Yes.
- E - Did you notice anything different in what I was saying to Fiona on the videotape?
- S - Yes, the second one, twice you passed a comment, I can't remember what it was. Something about that it was a very valuable answer to what the film was.
- E - O.K. Could you figure out why I was saying that?
- S - No, not really. Because some of the answers, I would have given different ones to what she had given.
- E - O.K. fine.

APPENDIX H

Complete List of Codings
for Personal and Empathetic Questions

| | Page |
|--|------|
| 1. <u>Coding of Responses to the Personal Question</u> | 218 |
| 2. <u>Coding of Responses to the Empathetic Question</u> | 219 |

Coding of Responses to the
Personal Question

| Sub. | Coding | I.R.C. | R.R.C. | Sub. | Coding | I.R.C. | R.R.C. |
|------|---------------|--------|--------|------|------------------|--------|--------|
| 1. | 4/6/1, 4/2, 5 | 4 | 4 | 32. | 2 | 2 | 2 |
| 2. | 2, 2 | 2 | 2 | 33. | 2 | 2 | 2 |
| 3. | 1/2 | 1 | 1 | 34. | 4/2/2, 1 | 4 | 2 |
| 4. | 2/1/1, 1/1, 1 | 2 | 1 | 35. | 2 | 2 | 2 |
| 5. | 1/2 | 1 | 1 | 36. | 1, 2/3, 1/2, 1/1 | | |
| 6. | 2/3, 2 | 3 | 3 | | 1/2, 1, 2, 1, 3 | 1 | 1 |
| 7. | 1/2 | 1 | 1 | 37. | 2, 2 | 2 | 2 |
| 8. | 5 | 5 | 5 | 38. | 2 | 2 | 2 |
| 9. | 1/2/2, 3/6 | 1 | 2 | 39. | 4/3/2 | 4 | 4 |
| 10. | 1 | 1 | 1 | 40. | 4/2 | 4 | 4 |
| 11. | 4/2 | 4 | 4 | 41. | 4/2 | 4 | 4 |
| 12. | 5, 1/3, 1 | 5 | 1 | 42. | 2 | 2 | 2 |
| 13. | 1/1/1 | 1 | 1 | 43. | 5 | 5 | 5 |
| 19. | 4/1 | 4 | 4 | 44. | 3/2 | 3 | 3 |
| 15. | 2 | 2 | 2 | 45. | 2/1/4 | 2 | 2 |
| 16. | 1, 2, 2/3, 2 | 1 | 2 | 46. | 2 | 2 | 2 |
| 17. | 2, 2 | 2 | 2 | 47. | 1 | 1 | 1 |
| 18. | 1, 5 | 1 | 1 | 48. | 1/2 | 1 | 1 |
| 19. | 1/2 | 1 | 1 | 49. | 2 | 2 | 2 |
| 20. | 4/2 | 4 | 4 | 50. | 1 | 1 | 1 |
| 21. | - | - | - | 51. | 2 | 2 | 2 |
| 22. | 2 | 2 | 2 | 52. | 6 | 6 | 6 |
| 23. | 2/2 | 2 | 2 | 53. | 1/2/2 | 1 | 2 |
| 24. | 2/2 | 2 | 2 | | | | |
| 25. | 2/2 | 2 | 2 | | | | |
| 26. | 1/2 | 1 | 1 | | | | |
| 27. | 1 | 1 | 1 | | | | |
| 28. | 4/2 | 4 | 4 | | | | |
| 29. | 2/2, 6 | 2 | 2 | | | | |
| 30. | 3, 4/2/2 | 3 | 3 | | | | |
| 31. | 1, 5 | 1 | 1 | | | | |

I.R.C. = Initial Response Category

R.R.C. = Representative Response Category

Coding of Responses to the
Empathetic Question

| Subject | Coding | Initial Response Category | Representative Response Category |
|---------|--------|---------------------------------|--|
| 1. | 4/4 | 4 | 4 |
| 4. | 4/1 | 1 | 1 |
| 6. | 4/2 | 2 | 2 |
| 7. | 4/2/2 | 2 | 2 |
| 8. | 5 | 5 | 5 |
| 10. | 2,2 | 2 | 2 |
| 11. | 4/5 | 5 | 5 |
| 14. | 4/1 | 1 | 1 |
| 15. | 2 | 2 | 2 |
| 18. | 2/4 | 2 | 2 |
| 19. | 4/2/2 | 2 | 2 |
| 28. | 1 | 1 | 1 |
| 29. | 4/6 | 6 | 6 |
| 36. | 4/1 | 1 | 1 |
| 37. | 4/2/2 | 2 | 2 |
| 38. | 1 | 1 | 1 |
| 45. | 5 | 5 | 5 |
| 50. | 2 | 2 | 2 |
| 52. | 4/1/1 | 1 | 1 |

APPENDIX I

Computer Programmes and Calculations

| | Page |
|--|------|
| 1. <u>Binomial Probabilities Programme</u> | 221 |
| 2. <u>Binomial Probabilities for N=18, p=.1667</u> | 222 |
| 3. <u>Binomial Probabilities for N=24, p=.1667</u> | 223 |
| 4. <u>Binomial Probabilities for N=30, p=.1667</u> | 224 |
| 5. <u>Fisher's Exact Test</u> | 225 |

Binomial Probabilities Programme

```
100 PRINT:PRINT"BINOMIAL PROBABILITIES PROGRAM"
120 PRINT:PRINT"N =";
130 INPUT N
140 PRINT"P =";
150 INPUT P
160 LET Q=1-P
170 PRINT:PRINT"BINOMIAL PROBABILITIES FOR N = "N", P ="P
180 PRINT:PRINT"K", "PROB(K)", "CUM.PROB", "INV.CUM.PROB"
190 LET F=0
200 FOR K=0 TO N
210 LET R=N-K
220 LET M=P^K*Q^R
230 LET C=1
240 FOR J=1 TO K
250 LET C=C*(N-J+1)/J
260 NEXT J
270 LET X=C*M
280 LET F=F+X
290 PRINT K,X,F,1-F
300 NEXT K
320 STCP
READY
```

Binomial Probabilities for N=18, p=.1667

| K | PROB(K) | CUM. PROB | INV. CUM. PROB | |
|----|--------------|-------------|----------------|---|
| 0 | .3756104E-1 | .3756104E-1 | .962439 | |
| 1 | .1352197 | .1727808 | .8272192 | |
| 2 | .2298736 | .4026543 | .5973457 | |
| 3 | .2451985 | .6478528 | .3521472 | |
| 4 | .1838988 | .8317516 | .1682484 | |
| 5 | .1029834 | .934735 | .6526502E-1 | α |
| 6 | .4462612E-1 | .9793611 | .2063891E-1 | |
| 7 | .1530038E-1 | .9946615 | .5338522E-2 | |
| 8 | .4207606E-2 | .9988691 | .1130916E-2 | |
| 9 | .9350234E-3 | .9998041 | .1958935E-3 | |
| 10 | .1683042E-3 | .9999724 | .275895E-4 | |
| 11 | .2448061E-4 | .9999969 | .3108755E-5 | |
| 12 | .2856072E-5 | .9999997 | .2533197E-6 | |
| 13 | .2636374E-6 | 1 | -.1024455E-7 | |
| 14 | .1883124E-7 | 1 | -.28871E-7 | |
| 15 | .1004333E-8 | 1 | -.2980232E-7 | |
| 16 | .3766248E-10 | 1 | -.2980232E-7 | |
| 17 | .8861761E-12 | 1 | -.2980232E-7 | |
| 18 | .9846401E-14 | 1 | -.2980232E-7 | |

Binomial Probabilities for $N=24$, $p=.1667$

| K | PROB(K) | CUM. PROB | INV. CUM. PROB |
|----|--------------|-------------|----------------|
| 0 | .1257911E-1 | .1257911E-1 | .9874209 |
| 1 | .6037975E-1 | .7295887E-1 | .9270411 |
| 2 | .1388734 | .2118323 | .7881677 |
| 3 | .203681 | .4155133 | .5844867 |
| 4 | .2138651 | .6293784 | .3706216 |
| 5 | .1710921 | .8004705 | .1995295 |
| 6 | .1083583 | .9088288 | .9117119E-1 |
| 7 | .5572714E-1 | .964556 | .3544405E-1 |
| 8 | .2368403E-1 | .98824 | .1176002E-1 |
| 9 | .842099E-2 | .996661 | .3339025E-2 |
| 10 | .2526297E-2 | .9991873 | .8127298E-3 |
| 11 | .6430574E-3 | .9998303 | .1696721E-3 |
| 12 | .1393291E-3 | .9999697 | .3034342E-4 |
| 13 | .257223E-4 | .9999954 | .4621223E-5 |
| 14 | .4042076E-5 | .9999994 | .5792826E-6 |
| 15 | .5389434E-6 | 1 | .4097819E-7 |
| 16 | .6063114E-7 | 1 | -.1955777E-7 |
| 17 | .570646E-8 | 1 | -.2514571E-7 |
| 18 | .4438358E-9 | 1 | -.2514571E-7 |
| 19 | .2803173E-10 | 1 | 1 |

-.2514571 * P

Binomial Probabilities for N=30, p=.1667

| K | PROB(K) | CUM. PROB | INV. CUM. PROB |
|----|--------------|-------------|----------------|
| 0 | .4212715E-2 | .4212715E-2 | .9957873 |
| 1 | .0252763 | .2948901E-1 | .970511 |
| 2 | .7330128E-1 | .1027903 | .8972097 |
| 3 | .1368291 | .2396194 | .7603806 |
| 4 | .1847193 | .4243387 | .5756613 |
| 5 | .1921081 | .6164468 | .3835532 |
| 6 | .1600902 | .776537 | .223463 |
| 7 | .1097761 | .8863131 | .1136869 |
| 8 | .6312129E-1 | .9494344 | .5056558E-1 |
| 9 | .0308593 | .9802937 | .1970628E-1 |
| 10 | .1296091E-1 | .9932546 | .6745365E-2 |
| 11 | .471306E-2 | .9979677 | .2032305E-2 |
| 12 | .1492469E-2 | .9994602 | .5398365E-3 |
| 13 | .4132993E-3 | .9998735 | .1265369E-3 |
| 14 | .1003727E-3 | .9999738 | .2616458E-4 |
| 15 | .2141285E-4 | .9999952 | .4752539E-5 |
| 16 | .401491E-5 | .9999993 | .7376075E-6 |
| 17 | .6612795E-6 | .9999999 | .7636845E-7 |
| 18 | .9551817E-7 | 1 | -.1862645E-7 |
| 19 | .1206546E-7 | 1 | -.2980232E-7 |
| 20 | .13272E-8 | 1 | -.3073364E-7 |
| 21 | .1264001E-9 | 1 | -.3073364E-7 |
| 22 | .1034183E-10 | | |

Fisher's Exact Test

```

100 REM/ FISHER'S EXACT TEST.  REQUIRES 2X2 CONTINGENCY TABLE.
101 REM/ ENTER LOWEST FREQUENCY FIRST.  MARGINALS ARE FIXED.
110 PRINT: PRINT" FISHER'S EXACT TEST."
120 PRINT: PRINT" FREQUENCIES":
130 INPUT A,B,C,D
140 LET N=A+B+C+D
150 REM/ ARRANGE CELLS SO A IS SMALLEST FREQ.  OF SMALLEST DIAG.
200 IF A*D<B*C GO TO 300
210 LET T1=A
220 LET T2=D
230 LET A=B
240 LET D=C
250 LET B=T1
260 LET C=T2
300 IF A<D GO TO 400
310 LET T1=A
320 LET T2=C
330 LET A=D
340 LET C=B
350 LET D=T1
360 LET B=T2
400 LET M1=A+B
410 LET M2=C+D
420 LET M3=A+C
430 LET M4=B+D
440 PRINT A ;B;M1
450 PRINT C ;D;M2
460 PRINT M3 ;M4
470 PRINT "N ="N
500 REM/ CALCULATE MOST EXTREME CASE
510 LET I=A
520 LET A=0
530 LET B=M1
540 LET C=M3
550 LET D=M2-C
560 PRINT: PRINT" EXTREME CASE"
570 PRINT A ;B;M1
580 PRINT C ;D;M2
590 PRINT M3 ;M4
600 REM/ CALCULATE P FOR EXTREME CASE
610 LET K=M2
620 GOSUB 900
630 LET F1=L
640 LET K=M4
650 GOSUB 900
660 LET F2=L
670 LET K=N
680 GOSUB 900
690 LET F3=L
700 LET K=D
710 GOSUB 900
720 LET F4=L

```

Fisher's Exact Test cont.

```
73Ø LET P=F1/F3*F2/F4
74Ø PRINT"P ="P
75Ø LET S=P
76Ø REM/ CALCULATE P FOR LESS EXTREME CASES.
77Ø FOR J=1 TO I
78Ø LET P=B/(A+1)*C/(D+1)*P
79Ø LET S=S+P
80Ø LET A=J
81Ø LET B=M1-A
82Ø LET C=M3-A
83Ø LET D=M2-C
84Ø PRINT:PRINT A;B;M1
85Ø PRINT C;D;M2
86Ø PRINT M3;M4
87Ø PRINT"P ="P
88Ø NEXT J
89Ø PRINT:PRINT"SUM P ="S
Ø99 STOP
90Ø LET L=1
91Ø FOR J=2 TO K
92Ø LET L=L*J
93Ø NEXT J
94Ø RETURN
```

BIBLIOGRAPHY

- Adams, R.S. & Biddle, B.J. Realities of teaching: Explorations with videotape. New York: Holt, Rinehart and Winstone, 1970.
- Allen, K.E., Henke, L.B., Harris, F.R., Baer, D.M., & Reynolds, N.J. Control of hyperactivity by social reinforcement of attending behaviour. In S.H. Osipow & W.B. Walsh (Eds.), Behavior change in counseling: Reading and cases. New York: Appleton-Century-Crofts. 1970, Pp.160-166
- Bandura, A. Aggression: A social learning analysis. Englewood Cliffs, New Jersey: Prentice Hall, 1973.
- Bandura, A. Influence of models' reinforcement contingencies on the acquisition of imitative responses. Journal of Personality and Social Psychology, 1965, 1, 589-595.
- Bandura, A. Principles of behavior modification. New York: Rinehart and Winston, 1969.
- Bandura, A. & Huston, A.C. Identification as a process of incidental learning. Journal of Abnormal and Social Psychology, 1961, 63, 311-318
- Bandura, A. & McDonald, F.J. Influence of social reinforcement and the behavior of models in shaping children's moral judgments. Journal of Abnormal and Social Psychology, 1963, 67, 274-281.
- Bandura, A., Ross, D., & Ross, S.A. A comparative test of the status envy, social power, and secondary reinforcement theories of identification learning. Journal of Abnormal and Social Psychology, 1963a, 67, 527-534.
- Bandura, A., Ross, D., & Ross, S.A. Imitation of film-mediated aggressive models. Journal of Abnormal and Social Psychology, 1963b, 66, 3-11.
- Bandura, A., Ross, D., & Ross, S.A. Vicarious reinforcement and imitative learning. In D.M. Gelfand (Ed.), Social learning in children. Belmont, California: Brooks/Cole, 1969, Pp. 196-207.
- Bandura, A. & Walters, R.H. Social learning and personality development. New York: Holt, Rinehart and Winston, 1963.
- Bartolone, P.I. Teachers objectives and questions in primary reading. The Reading Teacher, 1969, 23, 27-33.
- Binder, A., McConnell, D., & Sjolohn, N.A. Verbal conditioning as a function of experimenter characteristics. Journal of Abnormal and Social Psychology, 1957, 55, 309-314.

- Blalock, H.M., Jr. Social Statistics (2nd ed.) New York: McGraw Hill, 1972.
- Bloom, B.S. (Ed.) Taxonomy of educational objectives. London: Longmans-Green, 1956.
- Bloom, B.S. Hastings, J.T., & Madaus, G.F. Handbook on formative and summative evaluation of student learning. New York: McGraw-Hill, 1971.
- Bostrom, R.N., Vlandis, J.W., & Rosenbaum, M.E. Grades as reinforcing contingencies and attitude change. Journal of Educational Psychology, 1961, 52, 112-115.
- Brewer, J.K. On the power of statistical test in the American Educational Research Journal. American Educational Research Journal, 1972, 9, 391-401.
- Broden, M., Bruce, C., Mitchell, M.A., Carter, V., & Hall, R.V. Effects of teacher attention on attending behavior of two boys at adjacent desks. Journal of Applied Behavior Analysis, 1970, 3, 199-203.
- Cazden, C.D. Children's questions: Their forms, functions and role in education. Young Children, 1970, 25, 202-220.
- de Charms, R. & Rosenbaum, M.E. Status variables and matching behavior. Journal of Personality, 1960, 28, 492-502.
- Ditrichs, R., Simon, S., & Greene, B. Effects of vicarious scheduling on the verbal conditioning of hostility in children. Journal of Personality and Social Psychology, 1967, 6, 71-78.
- Dulany, D.E., Jr. Hypothesis and habits in verbal "operant conditioning." Journal of Abnormal and Social Psychology, 1961, 63, 251-263.
- Encyclopaedia Britannica Educational Corporation, Reflections on Time. Chicago: Encyclopaedia Britannica Educational Corporation, 1970.
- Epstein, R. Aggression toward outgroups as a function of authoritarianism and imitation of aggressive models. Journal of Personality and Social Psychology, 1966, 3, 574-579.
- Eriksen, C.W. Discrimination and learning without awareness: A methodological survey and evaluation. Psychological Review, 1960, 67, 279-300.
- Evans, G.W. & Oswalt, G.L. Acceleration of academic progress through the manipulation of peer influence. In C.E. Pitts (Ed.), Operant conditioning in the classroom, New York: Thomas Y. Crowell, 1971, Pp. 184-193.

- Ferster, C.B. & Perrott, M.C. Behavior principles. New York: Meredith, 1968.
- Galloway, C. Nonverbal communication. Theory into Practice, 1966, 7 (5), 172-175.
- Gelfand, D.M. The influence of self-esteem on rate of verbal conditioning and social matching behavior. Journal of Abnormal and Social Psychology, 1962, 65, 259-265.
- Glass, G.V. & Stanley, J.C. Statistical methods in education and psychology. Englewood Cliffs, New Jersey: Prentice Hall, 1970.
- Goetz, E.M. & Baer, D.M. Social control of form diversity and the emergence of new forms in children's block building. Journal of Applied Behavior Analysis, 1973, 6, 209-217.
- Greenspoon, J. The reinforcing effect of two spoken sounds on the frequency of two responses. American Journal of Psychology, 1955, 68, 409-416.
- Gronlund, N. Stating behavior objectives for classroom instruction. Toronto: Collier-Macmillan, 1970.
- Grusec, J.E. & Mischel, W. The model's characteristics as determinants of social learning. Journal of Personality and Social Psychology, 1966, 4, 211-215.
- Hall, R.V., Cristler, C., Cranston, S.S., & Tucker, B. Teachers and parents as researchers using multiple baseline designs. Journals of Applied Behavior Analysis, 1970, 3, 247-255.
- Hart, B.M., Reynolds, N.J., Baer, D.M., Brawley, E.R., & Harris, F.R. Effect of contingent and non-contingent social reinforcement on the cooperative play of a preschool child. In C.E. Pitts (Ed.), Operant conditioning in the classroom. New York: Thomas Y. Crowell, 1971, Pp. 294-300.
- Hart, B.M. & Risley, T.R. Establishing use of descriptive adjectives in the spontaneous speech of disadvantaged preschool children. Journal of Applied Behavior Analysis, 1968, 1, 109-120.
- Henderson, R.W. & Garcia, A.B. The effects of parent training program on the question-asking behavior of Mexican children. American Educational Research Journal, 1973, 10, 193-201.
- Hilgard, E.R. & Marquis, D.G. Conditioning and learning. (Rev. ed.) New York: Appleton-Century-Crofts, 1961.

- Hovland, C.I. Experimental studies in rote-learning theory, III: Distribution of practice with varying speeds of syllable presentation. Journal of Experimental Psychology, 1938, 23, 172-190.
- Huenecke, D. Cognitive levels of teacher objectives and oral classroom questions for curriculum guide users and non-users. Educational Leadership Supplement, 1970, 27, 378-383.
- Hunkins, F.P. Analysis and evaluation questions: Their effects upon critical thinking. Educational Leadership Research Supplement, 1970, 3, 697-705.
- Hunkins, F.P. Effects of analysis and evaluation questions on various levels of achievement. The Journal of Experimental Education, 1969, 38 (2), 45-58.
- Hunkins, F.P. Questioning strategies and techniques. Boston: Allyn and Bacon, 1972.
- Hurlock, E.B. An evaluation of certain incentives used in school work. The Journal of Educational Psychology, 1925, 16, 145-159.
- Jakubczak, L.F. & Walters, R.H. Suggestibility as dependency behavior. Journal of Abnormal and Social Psychology, 1959, 59, 102-107.
- Jensen, A.R. Spelling errors and the serial-position effect. Journal of Educational Psychology, 1962, 53, 105-109.
- Kagan, J. & Mussen, P.H. Dependency themes on the TAT and group conformity. Journal of Consulting Psychology, 1956, 20, 29-32.
- Kanareff, V.T. & Lanzetta, J.T. Effects of success-failure experiences and probability of reinforcement upon the acquisition and extinction of an imitative response. Psychological Reports, 1960a, 7, 151-166.
- Kanareff, V.T. & Lanzetta, J.T. Effects of task definition and probability of reinforcement upon the acquisition and extinction of imitative responses. Journal of Experimental Psychology, 1960b, 60, 340-348.
- Kanfer, F.H. Verbal conditioning: A review of its current status. In T.R. Dixon & D.L. Horton (Eds.), Verbal behavior and general behavior theory. Englewood Cliffs, New Jersey: Prentice-Hall, 1968, Pp. 254-290.
- Kanfer, F.H. & Marston, A.R. Human reinforcement: Vicarious and direct. Journal of Experimental Psychology, 1963, 65, 292-296.
- Kibler, R.J., Barker, L.L., & Miles, D.T. Behavioral objectives and instruction. Boston: Allyn and Bacon, 1970.

- Krasner, L., Weiss, R.L., & Ullman, L.P. Responsivity to verbal conditioning as a function of two different measures of "awareness." American Psychologist, 1959, 14, 388 (Abstract).
- Kriekhause, E.E. & Eriksen, C.W. A study of awareness and its effects on learning and generalizations. Journal of Personality, 1960, 28, 503-517.
- Krumboltz, J.D. & Schroeder, W.W. Promoting career planning through reinforcement. Personnel and Guidance Journal, 1965, 44, 19-26.
- Krumboltz, J.D. & Thoresen, C.E. The effect of behavior counseling in group and individual settings on information-seeking behavior. Journal of Counseling Psychology, 1964, 11, 324-333.
- Krumboltz, J.D. & Varenhorst, B.B., & Thoresen, C.E. Nonverbal factors in the effectiveness of models in counseling. Journal of Counseling Psychology, 1967, 14, 412-418.
- Lefkowitz, M.M., Blake, R.R., & Mauton, J.S. Status factors in pedestrian violation of traffic signals. Journal of Abnormal and Social Psychology, 1955, 51, 704-706.
- Lesser, G.S. & Abelson, R.P. Personality correlates of persuasibility in children. In C.I. Hovland & I.L. Janis (Eds.), Personality and persuasibility. New Haven: Yale University Press, 1959, Pp. 187-206.
- Levin, S. The effects of awareness on verbal conditioning. Journal of Experimental Psychology, 1961, 61, 67-75.
- Lovaas, O.I. A programme for the establishment of speech in psychotic children. In J.K. Wing (Ed.), Early childhood autism. Oxford: Pergamon, 1966, Pp. 115-144.
- Maccoby, E.E. & Wilson, W.C. Identification and observational learning from films. Journal of Abnormal and Social Psychology, 1957, 55, 76-87.
- MacMillan, D.L. Behavior modification in education. New York: Macmillan, 1973.
- Madaus, G.F., Woods, E.M., & Nuttall, R.L. A casual model analysis of Bloom's taxonomy. American Educational Research Journal, 1973, 10, 253-262.
- Marston, A.R. Variables in extinction following acquisition with vicarious reinforcement. Journal of Experimental Psychology, 1964, 68, 312-315.

- Marston, A.R. & Kanfer, F.H. Group size and number of vicarious reinforcements in verbal learning. Journal of Experimental Psychology, 1963, 65, 593-596.
- Matarazzo, J.D., Saslow, G., & Pareis, E.N. Verbal conditioning of two response classes: Some methodological considerations. Journal of Abnormal and Social Psychology, 1960, 61, 190-206.
- Mausner, B. Studies in social interaction, III: Effect of variation in one partner's prestige on the interaction of observer pairs. Journal of Applied Psychology, 1953, 37, 391-393.
- Mausner, B. The effect of one partner's success in a relevant task on the interaction of observed pairs. Journal of Abnormal and Social Psychology, 1954a, 49, 557-560.
- Mausner, B., & Bloch, B.L. A study of the additivity of variables affecting social interaction. Journal of Abnormal and Social Psychology, 1957, 54, 250-256.
- McDavid, J.W. Effects of ambiguity of environmental cues upon learning to imitate. Journal of Abnormal and Social Psychology, 1962, 65, 381-386.
- McDonald, F.J. Behavior modification in teacher education. In C.E. Thoresen (Ed.), Behavior modification in education. Chicago: University of Chicago Press, 1973, Pp.41-76.
- McGeoch, J.A. & Irion, A.L. The psychology of human learning. New York: David McKay, 1952.
- Mednick, S.A. Learning. Englewood Cliffs, New Jersey; Prentice-Hall, 1964.
- Miller, N.E. & Dollard, J. Social learning and imitation. London: Kegan Paul, Trench, Trubner, 1945.
- Mischel, W. & Grusec, J. Determinants of the rehearsal and transmission of neutral and aversive behaviors. Journal of Personality and Social Psychology, 1966, 3, 197-207.
- Neisworth, J.T. & Smith, R.M. Modifying retarded behavior. Boston: Houghton-Mifflin, 1973.
- Nuthall, G. Learning in classroom discussion: A theoretical explanation. Paper presented at the Annual Conference of the Australian and New Zealand Association of the Advancement of Science, Brisbane, 1971.
- O'Connor, R.D. Modification of social withdrawal through symbolic modeling. Journal of Applied Behavior Analysis, 1969, 2, 15-22.

- Odom, R.D., Leibert, R.M., & Hill, J.H. The effects of modeling cues, reward, and attentional set on the production of grammatical and ungrammatical syntactic constructions. Journal of Experimental Child Psychology, 1968, 6, 131-140.
- Peters, R.S. Education as initiation. In R.D. Archambault (Ed.), Philosophical analysis and education. Oxford; Routledge and Kegan Paul, 1965, Pp. 87-111.
- Peters, R.S. Ethics and education. Oxford: George Allen and Unwin, 1966.
- Peters, R.S. What is an educational process? In R.S. Peters (Ed.), The concept of education. London: Routledge and Kegan Paul, 1967, Pp. 1-23.
- Premack, D. Reinforcement theory. In D. Levine (Ed.), Nebraska symposium on motivation 1965. Lincoln, Nebraska: University of Nebraska Press, 1965, Pp. 123-180.
- Rosenbaum, M.E. & Bruning, J.L. Direct and vicarious experience of variation in percentage of reinforcement. Child Development, 1969, 37, 959-966.
- Rosenbaum, M.E. & Tucker, I.F. The competence of the model and the learning of imitation and nonimitation. Journal of Experimental Psychology, 1962, 63, 183-190.
- Rosenblith, J.F. Imitative color choices in kindergarten children. Child Development, 1961, 32, 211-223.
- Rosenblith, J.F. Learning by imitation in kindergarten children. Child Development, 1959, 30, 69-80.
- Rosenthal, T.L. & Whitebook, J.S. Incentives versus instructions in transmitting grammatical parameters with experimenter as model. Behavior Research and Therapy, 1970, 8, 189-196.
- Rosenthal, T.L., Zimmerman, B.J., & Durning, K. Observationally induced changes in children's interrogative classes. Journal of Personality and Social Psychology, 1970, 16, 681-688.
- Ross, D. Relationship between dependency, intentional learning, and incidental learning in preschool children. Journal of Personality and Social Psychology, 1964, 4, 374-381.
- Ryan, T.A. & Krumboltz, J.D. Effect of planned reinforcement counseling on client decision-making behavior. Journal of Counseling Psychology, 1964, 11, 315-323.
- Sanders, N.M. Classroom questions. New York: Harper and Row, 1966.

- Schein, E.H. The effect of reward on adult imitation behavior. Journal of Abnormal and Social Psychology, 1954, 49, 389-395.
- Scott, P.M., Burton, R.V., & Yarrow, M.R. Social reinforcement under natural conditions. In D.M. Gelfand (Ed.), Social learning in children. Belmont, California: Brooks/Cole, 1969, Pp. 245-254.
- Senter, R.J. Analysis of data. Glenview, Illinois: Scott, Foresman, 1969.
- Siegel, S. Nonparametric statistics for the behavioral sciences. N.Y.: McGraw Hill, 1956.
- Skinner, B.F. About behaviorism. New York, Alfred A. Knopf, 1974.
- Skinner, B.F. Contingencies of reinforcement. New York: Appleton-Century-Crofts, 1969.
- Skinner, B.F. Science and human behavior. New York: Macmillan, 1953.
- Smith, F.J. & Marston, A.R. Effects of inter- and intra-response class differences upon learning via vicarious reinforcement. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 360-364.
- Snow, R.E. Representative and quasi-representative designs for research on teaching. Review of Educational Research, 1974, 44, 265-291.
- Spielberger, C.D. & DeNike, L.D. Operant conditioning of plural nouns: A failure to replicate the Greenspoon effect. Psychological Reports, 1962, 11, 355-366.
- Staats, A.W. Learning, language and cognition. New York: Holt, Rinehart and Winston, 1968.
- Staats, A.W. & Butterfield, W. Treatment of nonreading in a culturally deprived juvenile delinquent: An application of reinforcement principles. Child Development, 1965, 36, 925-942.
- Suchman, J.R. The child and the inquiry process. In A.H. Passow & R.R. Leeper (Eds.), Intellectual development: Another look. Washington, D.C.: Association for Supervision and Curriculum Development, 1964, Pp. 59-77.
- Tanner, D. Using behavioral objectives in the classroom. New York: Macmillan, 1972.

- Tatz, S.J. Symbolic activity in "learning without awareness."
American Journal of Psychology, 1960, 73, 239-247.
- Travers, R.M.W., Reid, I.E., & Van Wagenen, R.K. Research on reinforcement and its implications for education.
Journal of Teacher Education, 1964, 15, 223-229.
- Travers, R.M.W. & Van Wagenen, R.K. Observational learning in a simulated recitation group. American Educational Research Journal, 1964, 1, 26-34.
- Trent, J.W. & Cohen, A.M. Research on teaching in higher education. In R.M.W. Travers (Ed.), Second handbook of research on teaching. Chicago: Rand McNally, 1973, Pp. 997-1071.
- Weinstein, W.K. & Lawson, R. The effect of experimentally-induced "awareness" upon performance in free-operant verbal conditioning and on subsequent tests of awareness.
The Journal of Psychology, 1963, 56, 203-211.
- Willemsen, E.W. Understanding statistical reasoning. San Francisco: W.H. Freeman, 1974.
- Zahorik, J.A. Teacher verbal feedback. Classroom Interaction Newsletter, 1969, 4, 34-41.
- Zimmerman, E.H. & Zimmerman, J. The alteration of behavior in a special classroom situation. In R.Ulrich, T. Stachnick, & J.Maby (Eds.), Control of human behavior. Glenview, Illinois: Scott, Foresman, 1966, Pp. 94-96.
- Zimmerman, E.H., Zimmerman, J., & Russell, C.D. Differential effects of token reinforcement of instruction-following.
Journal of Applied Behavior Analysis, 1969, 2, 101-112.