INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

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

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

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame.

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

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University Microfilms International
300 N. ZEEB ROAD, ANN ARBOR, MI 48106
18 BEDFORD ROW, LONDON WC1R 4EJ, ENGLAND
BERGAN, KATHRYN SUZANNE

THE EFFECTS OF PROBLEM EXEMPLAR VARIATIONS ON FRACTION IDENTIFICATION IN ELEMENTARY SCHOOL CHILDREN

The University of Arizona

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106

Ph.D. 1981
THE EFFECTS OF PROBLEM EXEMPLAR VARIATIONS
ON FRACTION IDENTIFICATION IN ELEMENTARY SCHOOL CHILDREN

by

Kathryn Suzanne Bergan

A Dissertation Submitted to the Faculty of the
DEPARTMENT OF EDUCATIONAL PSYCHOLOGY
In Partial Fulfillment of the Requirements
For the Degree of
DOCTOR OF PHILOSOPHY
In the Graduate College
THE UNIVERSITY OF ARIZONA

1981
THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Kathryn Suzanne Bergan entitled The Effects of Problem Exemplar Variations on Fraction Identification in Elementary School Children and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Dissertation Director

Date
STATEMENT BY AUTHOR

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

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

SIGNED: Kathryn S. Bogan
ACKNOWLEDGMENTS

The shaping of my life has been very complex. Many of the experiences playing a major role in its development have resulted from planned and happenstance, long and short encounters with others. To these people I owe a deep debt of gratitude and it is to them that I dedicate this dissertation. In particular I would like to thank my parents, whose early guidance was certainly of immense importance. I would also like to thank the numerous teachers and professors from whom I have learned along the way. Particularly notable in their assistance and guidance in recent years and in the development of this dissertation have been Dr. Anthony A. Cancelli and Drs. David C. Berliner, Dennis L. Clark, Thomas R. Kratochwill, Shitala P. Mishra and Glenn I. Nicholson. My deepest appreciation is reserved for my husband, Jack, and for our sons, John and David. It is their loving support, friendship and guidance that I treasure above all else.
TABLE OF CONTENTS

| LIST OF TABLES                          | vi  |
| LIST OF ILLUSTRATIONS                  | vii |
| ABSTRACT                               | viii|
| CHAPTER                                |     |
| 1. INTRODUCTION                        | 1   |
| Hypotheses and Questions of the Study  | 10  |
| Hypotheses for the Experimental        | 10  |
| Investigation                          |     |
| Questions for the Exploratory Verbal   | 12  |
| Report Study                           |     |
| 2. RELATED RESEARCH                    | 13  |
| The Social Learning Perspective        | 13  |
| Vicarious Rule Learning                | 14  |
| Rule Verbalization and Acquisition     | 18  |
| Modeling and Feedback                  | 20  |
| The Information Processing Perspective | 21  |
| Information Processing and the         | 22  |
| Structure of Knowledge                 |     |
| Information Processing Views on        | 30  |
| Learning                               |     |
| Information Processing Theory and      | 43  |
| Methods of Research                    |     |
| 3. METHOD                              | 50  |
| Subjects                               | 50  |
| Experimenters                          | 51  |
| Materials                              | 51  |
| Pretest                                | 52  |
| Prerequisite Skill Test                | 52  |
| Training Sheets                        | 53  |
| Posttests                              | 53  |
TABLE OF CONTENTS—Continued

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedures</td>
<td>54</td>
</tr>
<tr>
<td>Pretest</td>
<td>54</td>
</tr>
<tr>
<td>Prerequisite Group and Training Assignment</td>
<td>54</td>
</tr>
<tr>
<td>Training Group Variations</td>
<td>54</td>
</tr>
<tr>
<td>Posttest Problem Type Variations</td>
<td>56</td>
</tr>
<tr>
<td>Exploratory Phase</td>
<td>57</td>
</tr>
<tr>
<td>4. RESULTS</td>
<td>58</td>
</tr>
<tr>
<td>Modified Path Analysis Assessing</td>
<td></td>
</tr>
<tr>
<td>Training Effects</td>
<td>59</td>
</tr>
<tr>
<td>Error Analysis</td>
<td>74</td>
</tr>
<tr>
<td>Teacher Interview Analysis</td>
<td>78</td>
</tr>
<tr>
<td>Analysis of Think-Aloud Data</td>
<td>79</td>
</tr>
<tr>
<td>5. DISCUSSION</td>
<td>83</td>
</tr>
<tr>
<td>APPENDIX A: PRETESTING, TRAINING AND POSTTESTING MATERIALS</td>
<td>95</td>
</tr>
<tr>
<td>APPENDIX B: THINK-ALOUD PROTOCOLS</td>
<td>110</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>122</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Age, Grade, Sex and Ethnicity of Study Participants</td>
<td>50</td>
</tr>
<tr>
<td>2. $X^2$ Values for Eight Models of Independence</td>
<td>65</td>
</tr>
<tr>
<td>3. Observed Frequencies, Expected Frequencies, Odds and Odds Ratios for Training Groups and Problem Type Under the Hypothesis of No-Three-Way Interaction</td>
<td>72</td>
</tr>
<tr>
<td>4. Error Analysis for Fractions One-Third, One-Fifth, Two-Thirds and Two-Fifths. Hypothesis Evaluated: Mutual Independence of Training Group and Error Type</td>
<td>76</td>
</tr>
<tr>
<td>5. Observed Frequencies and Odds Ratios for Training Group and Error Type Data</td>
<td>77</td>
</tr>
</tbody>
</table>
LIST OF ILLUSTRATIONS

Figure                                                                 Page
1. Marginal Independence Contingency Table - Hypothesis: Joint Variable AB Independent of Variable C 63
2. Conditional Independence Contingency Table - Hypothesis: Variable B is Independent of Variable C within Levels of A 63
3. Path Diagram Describing Relations Between Variables A, B and C 71
4. First Two Posttest Responses Made by Think-Aloud Student Taught with One-Element Exemplars 82
5. First Two Posttest Responses Made by Think-Aloud Student Taught with Denominator-Rule Exemplars 82
ABSTRACT

A major purpose of this study was to investigate the relationship between fraction-identification rule abstraction and training exemplars. Fractions can be identified with at least two different rules. One of these, the denominator rule, is general in that it yields correct responses across a wide variety of fraction identification problems. The second, the one-element rule, is appropriate only when the number of elements in the set and the denominator-specified number of subsets are equivalent. Because these rules are not equally serviceable, a question of major importance is what factors determine which of the two fraction identification rules a child will learn during training.

The main hypothesis within this study specified that the nature of the fraction identification rule abstracted by a learner would be influenced by the nature of the examples used in training. It was further hypothesized that mastery of the denominator rule would positively affect performance on one-element problems, and that denominator-rule problem errors consistent with the one-element rule would occur significantly more frequently than would be expected by chance. The study addressed two additional questions. These related to recent work in the area of information
processing and concerned both changes in learner behavior across training/posttesting sessions and consistency between verbal reports of thinking processes and the fraction rules hypothesized to control fraction identification.

A pretest was used to determine eligibility to participate in the study. Eighty-two children incapable of set/subset fraction identification participated in the study. Two additional children were involved in an exploratory phase. The children ranged in age from six to ten and in grade level from one to four. The participants were mainly from middle and lower-class Anglo and Mexican-American homes. Children were randomly assigned to one of two training groups. In both training groups learners were provided, through symbolic (verbal) modeling, the general denominator rule for fraction identification. Children in one training group were also provided examples of fraction identification requiring the denominator rule. In the second training group children were provided simple examples in keeping with the denominator rule stated as part of instruction and yet also in keeping with the unstated one-element rule.

A modified path analysis procedure was used to assess the effects of training group assignment on fraction identification performance. Results of this analysis suggest a training group main effect. That is, children's performance on fraction identification posttest problems was in keeping with rules associated with the training examples they had
been provided. The results suggest that the strongest effects of training were related to performance on denominator rule problems in that the odds of passing the denominator rule posttest were 11.52 times greater for children taught with denominator-rule exemplars than for children taught with the one-element exemplars. The findings also suggest that performance on either of the fraction identification tasks influenced performance on the other. A further finding was that training with the one-element exemplars was associated with performance congruent with inappropriate use of the one-element rule. Recall that the one-element rule was never stated and that the exemplars, while compatible with the one-element rule were equally compatible with the stated denominator rule. Protocols from the children in the exploratory portion of the study suggest that the child taught with the ambiguous exemplars did abstract the one-element rule while the child taught with the denominator rule exemplars abstracted the denominator rule. The protocols also suggest that the child taught with the denominator rule made changes in his thinking as training and posttesting progressed.
CHAPTER 1

INTRODUCTION

An important aspect of human experience in general and of academic experience in particular is the learning of rules (Gagne, 1977; Rosenthal and Zimmerman, 1978). Rules make it possible to represent relations among stimuli thus allowing individuals to respond in a consistent manner to classes of stimuli. Academic functioning is affected by rules in a variety of ways. For example, the field of mathematics is composed of complex rule systems. Learning the fundamental rules making up these systems is essential not only for adequate understanding of mathematical processes, but also for effective performance in the solving of mathematical problems.

A number of theorists (e.g., Bandura, 1977; Brown and Burton, 1978; Gagne, 1962, 1970, 1977; Newell and Simon 1972) have been concerned with the role that rules play in intellectual functioning and have studied the conditions leading to rule learning. During recent years social learning theorists have taken a prime role in advancing knowledge with respect to the manner in which rules are learned. Social learning theory stresses the role of vicarious learning in rule acquisition. Social learning theorists take the
position that when a learner observes stimuli reflecting a rule, the learner will tend to acquire the rule. For example, if a child observes a teacher using one or more rules to solve a set of mathematical problems, it is assumed that the child will tend not only to learn to solve the particular problem modeled but also how to solve other problems within the same classification, that is, within the category of problems solvable with the rules demonstrated by the teacher (Rosenthal and Zimmerman, 1978). Social learning theorists do not require rule statement as proof of rule acquisition. Rather, the ability to generalize beyond the specific problems taught is taken as an indication of rule acquisition.

The general paradigm employed to study rule acquisition within the social learning framework is to expose one or more randomly selected groups to training conditions in which the rule under examination is modeled. The performance of the experimental group(s) is generally compared to the performance of a control group under the assumption that the experimental group(s) will exhibit significantly more rule-consistent behavior than the control group. The use of this paradigm has led to data suggesting that a child can, through the observation of rule-consistent behavior, acquire patterns which span a wide diversity of specific tasks (Zimmerman and Rosenthal, 1974b). These acquired patterns have been shown to be quite independent of the specific
modeled experience, that is, they do not suggest mimicry. Learners have been able to respond consistently with the inferred governing rule when presented a new stimulus arrangement or entirely new content (Carroll, Rosenthal and Brysh, 1972; Rosenthal and Whitebook, 1970; Zimmerman and Rosenthal, 1974b).

While social learning theorists have focused on the processes involved in rule acquisition, information processing theorists (e.g., Brown and Burton, 1978; Newell and Simon, 1972) have looked at the manner in which rules are used in information processing. Much of the work in information processing has focused on problem solving. One particularly important insight to come out of information processing research on problem solving is the finding that human problem solvers often employ defective or incomplete rules in problem solving. For example, Brown and Burton (1978) found that children manifested a variety of defective rules (which Brown and Burton called bugs) in solving arithmetic problems. Several other investigators have attained similar findings. For instance, Carry, Lewis and Bernard (1978) observed that students made a variety of errors that reflected the application of defective rules when solving algebra problems.

The widespread occurrence of defective rules observed in information processing studies suggests that in the course of rule learning learners may often acquire rules
other than those targeted for instruction. In the past it has often been implicitly assumed that when learners fail to acquire rules through instruction, they learn nothing which is of significant interest. This assumption is reflected both in the pass-fail scoring procedures widely used to assess learner performance in academic settings and in the general research paradigm which has been employed by social learning theorists in the study of rule acquisition. In the social learning research paradigm, the emphasis has been placed on between group differences in amount of rule-consistent behavior and has not been focused on the non-rule-concordant behavior emitted by either experimental or control group members. Both the assessment and the research procedures treat incorrect performance as involving random error (Carry et al., 1978) arising from lack of knowledge or skill or from various causes such as carelessness. Information processing theory suggests that in many cases defective performance may reflect the application of defective rules rather than the total absence of knowledge or carelessness.

An unanswered question of theoretical and practical importance raised by the information processing findings is that of determining the conditions in the environment and in the learner that foster the acquisition of defective rules (Brown and Burton, 1978). One environmental factor that may promote defective rule acquisition has to do with ambiguity
associated with the exemplars used in instruction to facilitate rule learning. For instance, it may happen that a learner will be exposed to exemplars that reflect more than one rule. Under these circumstances the learner may acquire a rule that will produce correct performance in some cases, but not in others. Brown and Burton (1978) give an example involving a rule of this kind. They point out that the subtraction performance of some children suggests that the children follow a simple rule which holds that in subtraction one should always subtract the smaller number from the larger number. This rule works quite well until a problem requiring regrouping is encountered. On regrouping problems children following the "smaller from larger" rule will avoid borrowing by subtracting the number in the minuend from the number in the subtrahend. This approach obviously will yield an incorrect answer.

Exposure to ambiguous exemplars may come about inadvertently as teachers attempt to provide their students with simple, easy to comprehend examples of rules to be learned. It is unfortunately the case that simple examples may often be reflective of both a complex rule and a simple, less general rule. Consider a paper and pencil task often used in fraction identification instruction. In this task the teacher may demonstrate the fractional part of a set by placing Xs on a portion of the elements, for example, on two circles out of five to represent the fraction two-fifths.
In a study of fraction identification behavior, Bergan, Towstopiat, Cancelli and Karp (1981) assert that exemplars of the type just described are consistent with two rules. One of these rules, referred to as the denominator rule by Bergan and his associates, states that "in identifying a fraction with a denominator value of \( r \) for a set of \( n \) objects, the set of \( n \) objects must be partitioned into \( r \) equivalent subsets." The second rule, called the one-element rule by Bergan et al. equates subset and element within the set of \( n \) objects. This rule is less complex than the denominator rule as it does not include the set partitioning step. Set division into equivalent subsets has been omitted by considering each element of the set as a subset. A person using the one-element rule might successfully identify two-fifths of five objects by attending to the two in the numerator of the fraction and placing an X on each of two objects. This approach can completely avoid partitioning of the set into five equal subsets.

A teacher demonstrating fraction identification using exemplars which can be correctly solved with both the one-element and the denominator rule opens the door for the acquisition of either rule. Which rule will be acquired? This is an important question because the answer to it may shed some light on the manner in which defective rules are learned. A major purpose of the present research is to
investigate rule acquisition occurring as a function of the modeling of ambiguous fraction identification exemplars.

One reasonable hypothesis that may be advanced with respect to rule acquisition involving ambiguous exemplars is that when rule consistent behavior is modeled with exemplars consistent with both a simple and a complex rule, the simple rule will tend to be acquired. This hypothesis would suggest that children exposed to fraction identification exemplars consistent with both the one-element and the denominator rules would tend to acquire the one-element rule. If this were the case, the children would be expected to perform well on fraction identification problems consistent with the one-element rule. However, their performance would be poor on tasks involving exemplars requiring application of the denominator rule. Moreover, their errors on such tasks ought to tend to be consistent with the one-element rule. For example, children using the one-element rule would be expected to have difficulty identifying two-fifths of ten objects since correct identification of two-fifths of ten requires that the set of ten be partitioned into five equal subsets each containing two elements. Furthermore, it would be reasonable to expect some tendency for such children to err by placing Xs on two objects rather than four since that mistake would be consistent with the one-element rule.
Given the assumption that the use of ambiguous exemplars promotes defective rule acquisition, it is reasonable to assume that flawed rule learning could be avoided by using unambiguous exemplars. This would suggest that in the case of fraction identification children exposed to exemplars which can be correctly solved only through the use of the denominator rule ought to acquire that rule. For instance, if children were exposed to exemplars such as two-fifths of ten, it would be reasonable to expect that there would be some tendency for them to acquire the denominator rule. Rule acquisition could be inferred from accurate performance both on simple problems consistent with the one-element rule and the denominator rule and on more complex problems consistent only with the denominator rule.

Defective rule acquisition may be influenced not only by environmental conditions such as the types of exemplars modeled for learners, but also by conditions within the learner. Robert Gagne (1962, 1970, 1977) has suggested that one important condition that may influence rule acquisition is the extent to which learners possess prerequisite competencies associated with the rule learning tasks for which instruction is being provided. Gagne's view implies that children who initially lack prerequisites must either acquire them during the course of instruction or fail to master the tasks being taught. Burke, Hiber and Romberg (1977) have identified a number of prerequisites to
fraction identification. Gagne's notions suggest that these must be present in order for the learner to master fraction identification tasks. For instance, Burke and his associates assert that the ability to partition a set into subsets of equal size is prerequisite to the identification of fractions of the type requiring application of the denominator rule. Gagne's theory suggests that partitioning skill must be present if learners are to master fraction identification tasks calling for application of the denominator rule.

What Gagne's theory does not specify is what children will learn when they lack the necessary prerequisites to profit from instruction in the manner intended. One hypothesis that may be considered is that they will tend to acquire rules consistent with their present competencies. In the case of fraction identification, this would suggest that there would be some tendency for children lacking the partitioning prerequisite to acquire the one-element rule as a result of fraction identification instruction. This suggestion is based on the fact that the one-element rule does not require set partitioning.

The study will examine the research questions discussed in this chapter through the testing of formal hypotheses listed below. In addition, the investigation will include an exploratory phase in which verbal reports of the problem solving of a small number of subjects will be
studied intensively. It is hoped that the verbal report data will not only corroborate the assumptions examined through formal hypothesis testing but also may add new insights into the manner in which children solve fraction identification problems.

Hypotheses and Questions of the Study

The hypotheses to be examined in the study involve fraction-identification problems that vary with respect to the types of rules required for correct performance. Problems in which the denominator subsets are composed of a single element and are thus consistent with both the one-element and the denominator rules will be called one-element problems. Problems in which denominator subsets are composed of more than one element and are thus consistent with only the denominator rule will be called denominator-rule problems. In training examples, denominator-rule problems will consist of subsets containing three elements while in posttesting examples denominator-rule problems will have two-element subsets.

Hypotheses for the Experimental Investigation

Hypothesis 1—Children exposed to the denominator-rule exemplars during training will perform significantly better on denominator-rule problems but not on one-element problems than will children exposed to the one-element training exemplars.
Hypothesis 2—Mastery of the denominator-rule will positively affect performance on one-element problems.

Hypothesis 3—Irrespective of their subset division capability, children exposed to one-element training exemplars will not perform as well as children who are capable of subset division and are exposed to denominator-rule exemplars when tested on denominator-rule problems.

Hypothesis 4—Children with prerequisite skills will perform significantly better than children without prerequisite skills on problems in which subsets contain more than one element but not on problems in which subsets contain a single element.

Hypothesis 5—For problems containing subsets with more than one element, a significant interaction between exemplar type and prerequisite-ness will obtain while for problems with one-element subsets an exemplar main effect will occur.

Hypothesis 6—Denominator-rule problem errors consistent with the one-element rule will occur significantly more frequently than would be expected by chance.

Hypothesis 7—The proportion of denominator-rule problem errors compatible with the one-element rule should be significantly greater among children incapable of subset division than among children capable of subset division.
Questions for the Exploratory Verbal Report Study

**Question 1**—Will verbal report indicate thinking processes consistent with the one-element and denominator rules hypothesized to control fraction-identification performance in the present study?

**Question 2**—Will verbal reports of children obtained during learning and posttesting reveal attempts to increase efficiency in the fraction-identification process by using fewer steps than those demonstrated.
CHAPTER 2

RELATED RESEARCH

The present research is based on ideas coming from two different theoretical perspectives: social learning theory and information processing theory. These perspectives are relevant to the present research in two ways. First, these orientations have influenced the formation of the hypotheses to be examined. Their second influence is related to the choice of methods to be used to examine the cognitive processes under consideration. The present chapter reviews the literature on social learning and information processing perspectives with respect to both the hypotheses to be examined and the methods used to study them.

The Social Learning Perspective

A major focus of contemporary social learning research related to cognitive processes has been on rule-learning. Social learning theorists define a rule as a systematic relationship among objects or events (Rosenthal and Zimmerman, 1978) and consider rule-governed behavior to be an example of complex cognitive functioning. From the contemporary social learning perspective, performance on mathematical tasks of the type being investigated in the present study lies within the category of rule-governed behavior.
For example, in learning to identify a fractional portion of a set of objects, rules must be formed to guide fraction-identification behavior. Children engaged in this undertaking may learn that when they are shown a denominator of five they should partition the overall set of objects into five parts. The rule in this example relates an object, the denominator of five, with an event, set partitioning. The rule-governed behavior is rule-consistent performance on occasions appropriate to the use of the rule.

Social learning theorists have conducted a considerable amount of research identifying the factors influencing rule acquisition. This research provides the basis not only for certain hypotheses to be examined in the present study but also for constructing the instructional methods to be used in the investigation.

Vicarious Rule Learning

A fundamental premise upon which the present research is based is the social learning view that learners can acquire rules as a result of observing the behavior of a model. The assumption of vicarious rule learning forms the cornerstone for the instructional procedures to be used in the current investigation in that modeling will be employed in all training conditions.

Research on vicarious learning forms the basis of modern social learning theory. Contemporary social learning
theorists (e.g., Bandura 1977; Mischel, 1973; Rosenthal and Zimmerman, 1978) hold the view that learning can occur in the absence of reinforcement of overt behavior. This position on reinforcement differentiates contemporary social learning theory from earlier versions of the social learning position (e.g., Dollard and Miller, 1950; Miller and Dollard, 1941) and from the radical behaviorist views of Skinner (1953). Both the early social learning and radical behaviorist perspectives take the view that learning involving the imitation of modeled behavior requires reinforcement of overt responses that match the model's behavior (Dollard and Miller, 1950, Miller and Dollard, 1941; Skinner, 1953). By contrast, contemporary learning theorists under the leadership of Albert Bandura (1969, 1977) have taken the position that reinforcement is not necessary for observational learning. Following Tolman (1948), advocates of the contemporary social learning view make a distinction between learning and performance. They assert that reinforcement does not play a role in learning but rather serves an incentive function that can influence performance of learned responses.

Bandura (1965) conducted a highly influential study in support of the contemporary social learning view. Bandura exposed three groups of children to a set of modeling stimuli. One group observed the model being punished for his behavior. The second group saw the model being ignored for his actions and the third group observed the
model being rewarded for his behaviors. The group that observed the rewarded model showed the greatest amount of imitative behavior. However, in the absence of further training, all of the children were subsequently promised rewards if they would demonstrate what they had observed. Under these conditions, all the children displayed comparable levels of learning.

Although Bandura was able to demonstrate vicarious learning, some theorists assumed that the kind of learning that could occur through vicarious processes would be highly limited in scope. More specifically, it was argued that observational learning was a form of mimicry that could not explain the acquisition of complex cognitive competencies (Rosenthal and Zimmerman, 1978). Social learning theorists conducted several studies to refute this view. For example, Rosenthal, Zimmerman and Durning (1970) showed that children could acquire rules governing their question-asking behavior as a result of observing question-asking behavior of a model. In this investigation, the model showed a picture to the child and then asked a question associated with the picture. Different groups of children observed questions representing different rule categories. For example, one category included questions about the perceptual characteristics of stimuli, such as "What shape is that?". Another category involved casual relations illustrated by questions such as "How come the guitar makes music?". Children
in the study imitated the categories of question-asking behavior to which they were exposed rather than specific questions that had been modeled.

Observational rule learning was demonstrated in numerous other studies also. For example, children's sentence construction structures were shown to be influenced by modeling experiences as were their use of differing verb tenses (Bandura and Harris, 1966; Carroll et al., 1972; Liebert, Odom, Hill and Huff, 1969; Odom, Liebert and Hill, 1968; Rosenthal and Carroll, 1972). Other investigators used modeling procedures to teach rules related to moral judgement (Bandura and McDonald, 1963), equivalence concepts (Rosenthal, Moore, Dorfman and Nelson, 1971), and seriation (Bergan and Jeske, 1980; Henderson, Swanson and Zimmerman, 1975; Jeske, 1978; Swanson, 1976).

The social learning view that rules can be acquired through observation of a model did not go unchallenged. A serious criticism of social learning research on rule acquisition involved the assumption that modeling cued the performance of rules already in the repertoire of the observer. According to this position, observers actually learn nothing new as a result of exposure to a model. The no-new-learning criticism was particularly prevalent among Piagetians (e.g., Turiel, 1973) who followed the Piagetian (1962) position that imitation is limited by a child's level of intellectual development.
Social learning theorists conducted several investigations to demonstrate that new rule learning can occur as a result of exposure to modeled behavior (Zimmerman and Rosenthal, 1972; Zimmerman and Rosenthal, 1974a). For example, Zimmerman and Rosenthal (1974a) used modeling to teach children conservation skills that Piaget had argued could be learned only through discovery. In this experiment a model demonstrated correct judgments on various conservation tasks and supported these with appropriate verbal explanations. Children exposed to modeling conditions out-performed children in a no-instruction control group. Moreover, the modeling effects persisted on a retention test involving items on which the subjects were not trained.

Rule Verbalization and Acquisition

Modeling need not take the form of an overt physical act to influence rule learning. Symbolic modeling in the form of rule verbalization may also affect learning. For example, in teaching fraction identification a teacher might illustrate set partitioning by saying: "Now I am going to divide the set into five equal parts. To do this I will draw a circle around each of the five objects in the set."

Social learning theorists have found that symbolic modeling in the form of rule verbalization can have a beneficial influence on rule acquisition. For example, Zimmerman and Rosenthal (1972) examined the role of rule
verbalization and overt behavior modeling in the acquisition of a novel rule. Modeling accompanied by rule verbalization produced the highest level of rule learning and generalization. Rule verbalization in the absence of behavioral modeling also had a significantly beneficial influence on rule learning. The pattern of results found during acquisition and generalization was maintained during a retention task.

Further support for the roles of symbolic and behavioral modeling in skill acquisition is to be found in a fraction identification study by Karp (1978). The rule learners in Karp's investigation were exposed to one of four conditions: behavioral modeling, symbolic modeling, behavioral plus symbolic modeling and no modeling. Behavioral modeling in this investigation consisted of a model performing the fraction identification task; symbolic modeling consisted of a statement of the specific rule as it related to each fraction identification task. As in the Zimmerman and Rosenthal (1972) investigation, children in Karp's study receiving behavioral plus symbolic modeling performed significantly better than children receiving no modeling or than those receiving only symbolic modeling. However, unlike the results obtained by Zimmerman and Rosenthal (1972), Karp's data suggested that behavioral modeling alone was significantly better than symbolic modeling alone in the facilitation of rule learning and was not significantly different from combined behavioral/symbolic modeling effect.
Modeling and Feedback

Kulhavy (1977) has defined feedback as any one of a large group of techniques that can be used to relay to a learner information concerning the correctness/incorrectness of his/her task performance. One form of feedback is to provide the learner with an evaluative statement, such as yes or no, following performance. Zimmerman and Rosenthal (1974b) conducted a study designed to investigate the role of evaluative feedback in a concept attainment task. The learners, Chicano and Anglo youngsters, were exposed to symbolic modeling treatments and then were asked to make stimulus selections based on the rules which had been modeled. Evaluative feedback statements were given following stimulus selections. Data analysis revealed that both symbolic and behavioral modeling treatments created significant acquisition and retention effects while feedback played a rather minor role but did contribute a significant effect.

Bergan and Parra (1979) investigated the effect of behavioral modeling feedback on letter learning in Anglo and bilingual Mexican-American children. These researchers found that feedback consisting of both an evaluative statement and symbolic modeling was more effective than behavioral modeling alone but less effective than a combination of feedback with symbolic modeling and behavioral modeling. Bergan and Jeske (1980) in a study of the relationships among modeling, feedback, and the learning of seriation hierarchy
tasks obtained similar results in that they found feedback consisting of behavioral modeling to be more effective than the combination of guided practice and symbolic modeling in facilitating transfer for one of their tasks. However, for the remaining tasks Bergan and Jeske found that experimental condition did not contribute significantly to skill acquisition or transfer.

A study by Swanson (1976) provides further evidence related to the facilitative role of feedback. In her study she investigated the effects of seriation task training on transitivity task performance and the effects of varying instruction conditions. Her data suggest that both modeling plus the opportunity to practice and modeling combined with evaluative feedback were associated with greater learning than was modeling alone.

The Information Processing Perspective

Information processing theory conceives of a human being engaged in cognitive activity as an information processing system (IPS) (Newell and Simon, 1972). The basic components of an IPS include receptors that take in information from the external environment and from stimulus sources operating from within the system; effectors that can act upon the environment, a processor that interprets and in other ways operates on information in the system, and a memory that holds information in storage. Information processing
theory contributes to the present research in three ways. First, it provides means to represent the structure or organization of the knowledge associated with fraction identification. Second, it affords procedures to represent cognitive learning, and third, it involves innovations in method for studying cognitive phenomena that have been incorporated into the exploratory phase of the present research.

Information Processing and the Structure of Knowledge

Determining the structure or organization of knowledge is a major focus of information processing theory (Greeno, 1980). Information processing theorists have represented knowledge structures in a number of different ways. The formalisms used in the information processing approach to represent structure provide the basis in the present study for hypotheses about the manner in which children solve fraction identification problems and about the underlying skills required to bring about successful solutions to the problems.

The Hierarchical Structure of Knowledge—One of the most widely used formalisms to represent knowledge structure in educational circles is Robert Gagne's learning hierarchy model. Gagne (1962) set forth the view that complex cognitive skills arise through the mastery of a set of hierarchically organized subskills. At the time that he introduced his hierarchical model, Gagne did not conceptualize
hierarchies in information processing terms. However, he has since linked his perspective to the information processing viewpoint (Gagne, 1977).

There are two basic hypotheses associated with the learning hierarchy model (Bergan, 1980; White and Gagne, 1974). One is that acquisition of a superordinate skill requires mastery of hierarchically related subordinate skills. The second is that mastery of subordinate skills will have a beneficial influence on the acquisition of superordinate competencies.

Gagne's views have had a marked impact on educational research in general and on studies related to fraction concepts in particular. Burke et al. (1977) conducted an extensive survey of research and theory on fraction identification. Following their survey, these authors composed a working paper the main purpose of which was to present a careful re-examination of the initial instruction of fractions. As part of the working paper, Burke, Hiber and Romberg present an initial fraction hierarchy incorporating the works of Novillis (1976) and Harvey, Green and McLeod (1975) as well as the Initial Fraction Sequence developed at the University of Michigan. Within the hierarchy developed by Burke et al. several skills are presented as prerequisites to fraction identification as it is being defined within the current investigation. Included as initial prerequisites are the skills of counting and assessing
equivalence of groups. An additional subordinate skill to the skill of fraction identification with multiple-element subsets is the skill of equivalence partitioning. Equivalence partitioning is defined as the division of a set of elements into a specified number of subsets, each containing the same number of elements as the others.

Bergan et al. (1981) conducted hierarchy research on fraction identification that bears directly on the present study. These investigators studied children's performance on fraction identification tasks involving specification of the fractional portion of a set of objects. Bergan and his colleagues hypothesized that tasks that could be performed correctly through procedures consistent with the one-element rule described in the preceding chapter would be prerequisite to tasks requiring application of the denominator rule. For example, they assumed that the identification of two-fifths of five objects would be subordinate to the identification of two-fifths of ten objects. Their results supported this assumption.

Bergan and his associates recognized that children engaged in fraction identification would not necessarily be able to state either the one-element rule or the denominator rule in verbal form. They asserted simply that the behavior of the children would be consistent with these rules. This leaves open the question of what children actually do when they solve fraction identification problems. For example,
Bergan and his colleagues suggest that children whose performance is consistent with the one-element rule may identify fractions simply by emitting responses consistent with the fraction numerator. Thus, in identifying the fraction two-thirds, the children might put Xs on two objects.

Novillis (1976) reported findings consistent with those of Bergan and his colleagues. She also found that recognition of fractions such as two-fifths of five was prerequisite to recognition involving fractions such as two-fifths of ten. Unfortunately, Novillis used an early validation technique employed in some of the first hierarchy studies by Gagne and his associates. This technique does not provide an adequate test of hypothesized prerequisite relations (White, 1973). Nonetheless, Novillis' findings do provide additional support for the results obtained in the Bergan et al. (1981) study.

The results of the Bergan et al. and Novillis investigations indicate that some children emit behavior consistent with the one-element rule. This finding leads directly to the major question posed in the present study, namely, how is behavior consistent with the one-element rule acquired?

Although Gagne's learning hierarchy model has provided a useful framework for investigating knowledge structure, there are limitations to the approach. The learning model analyzes task performance in terms of hierarchically
organized subtasks. It is also important to be able to determine the sequence of steps that individuals follow in performing a single task. Much of the work in information processing theory on knowledge structures addresses the question of structure associated with the performance of a single task or class of tasks.

The Production System—One of the most widely used formalisms for describing a knowledge structure associated with the performance of a single task is the production system (Newell and Simon, 1972; Simon, 1979). The concept of the production system was advanced to describe the manner in which sequences of action are controlled in problem solving (Simon, 1979). A production system is composed of a set of instructions called productions. Each production is comprised of two components, a condition and an action. The basic rule governing the operation of the system is that whenever the condition associated with a particular production is present, the action implied by the production will be carried out. For example, in fraction identification, a production might be constructed for partitioning a set into subsets of equal size in accordance with the value of the denominator of the fraction. In this example, the denominator value could be regarded as the condition in the production. The process of set partitioning would be the action associated with this condition.
Production systems have been widely used in research on problem-solving competence (e.g., Newell and Simon, 1972) and in studies of learning during problem solving (Anderson, 1980; Anzai and Simon, 1979; Larkin, 1979). Nonetheless, production systems have certain noteworthy limitations. One is that in certain cases they may not make adequate provisions for consideration of strategic knowledge of the type involved in overall planning in problem solving (Greeno, 1978). The second is that they generally do not include provisions for the consideration of error in problem-solving performance (Brown and Burton, 1978).

**Procedural Networks**—The procedural network is a formalism that has been used increasingly during recent years to represent knowledge structures involving planning. The procedural network was introduced by Sacerdoti (1977). Procedural networks include condition/action relations similar to those involved in a production system. However, in a procedural network actions are typically associated not only with preconditions but also with consequences and with other component actions. Consequently, the coordination of actions must occur in accordance with the sequencing with which conditions are introduced and tested. The procedural network makes it possible to use planning mechanisms in problem solving that consider consequences of action in relation to preconditions of actions that may not be performed until much later in the problem-solving process.
Brown and Burton (1978) developed a procedural network for solving subtraction problems that is particularly relevant to the present research. The Brown and Burton network is based on the view that models of the structure of mathematical problem-solving behavior can profit from the inclusion of learner's misconceptions or errors as part of the structure of knowledge. Brown and Burton use the term "bugs" to describe errors occurring during the course of problem solving. Bugs are not careless errors; rather the term refers to systematic misconceptions associated with problem-solving activity.

Brown and Burton describe problem solving using the psycholinguistic concepts (Chomsky, 1968) of deep structure and surface structure. They use the term surface structure to refer to the actual record of problem-solving performance, for example, written performance on subtraction tasks. The deep structure indicates the underlying knowledge structure responsible for overt performance. The procedural network is used to represent this underlying structure.

In developing procedural networks for representing misconceptions as variants of correct performance, Brown and Burton (1978) began by constructing a network reflecting accurate problem-solving behavior. Their network consists of a collection of procedures with annotations in which the mechanism for sequencing procedure activation are made
explicit. Each procedure includes two main parts, a conceptual part indicating the intent of the procedure and an operational part specifying the method used to carry out the procedure. For example, the intent of the procedure might be to subtract one number from another. The methods used to accomplish subtraction might include standard procedures involving such techniques as borrowing or might be based on counting backwards from the number from which the subtraction is to be made.

Brown and Burton's (1978) use of the term procedural network differs somewhat from the manner in which Sacerdoti (1977) employed the concept. More specifically, the Brown and Burton procedural network does not incorporate extensive mechanisms for planning. The subtraction process is basically algorithmic in nature (Brown and Burton, 1978). That is, little strategic knowledge or planning is required in carrying out the subtraction process. It is worth noting that fraction identification should share the algorithmic character of the subtraction process.

The central importance of Brown and Burton's ideas for the present research lies in recognition of the fact that knowledge structures can usefully incorporate information about errors in thinking as well as information regarding accurate performance.
Information Processing Views on Learning

During recent years a number of information theorists have turned their attention to the study of learning. Research on learning has focused on computer simulation of learning occurring during the course of problem-solving activities. The central mechanism used to model the learning process has been the production system. Production systems applied in computer simulations of learning are called adaptive (Simon, 1979). Systems of this kind are able to build new productions and add them to the original system.

Learning and the Adaptive Production System--Anzai and Simon (1979) constructed an adaptive production system describing mechanisms by which learning may occur during the course of problem-solving. These investigators simulated problem solving using the "Tower of Hanoi" puzzle, a problem-solving task widely used in information processing research (Simon, 1979). The puzzle as used in the Anzai and Simon study involved three pegs placed on a board and labeled from left to right, A, B, C. Five discs varying in size were placed on peg A in ascending order with the largest disc on the bottom. The task was to transfer the discs to peg C while adhering to specified rules. The rules included a prohibition against placing any disc on another that was smaller than itself and a prohibition against moving a disc that had a disc on top of it.
Anzai and Simon began their work through a detailed analysis of the learning of a single human problem solver. As this subject engaged in repeated attempts to solve the Tower of Hanoi problem, she was instructed to reveal her problem-solving procedures by thinking aloud. She very quickly learned to solve the puzzle. However, repeated presentations of the problem eventuated in changes in her approach to the task.

Anzai and Simon identified four strategies that evolved from successive attempts to solve the Tower of Hanoi puzzle. The first they called the selective search strategy. This strategy involved the selection of legal moves and the elimination of illegal moves. Application of this simple strategy markedly reduced the size of the search space representing the possible moves that might be employed in reaching problem solution. Soon after the subject developed an effective search strategy, she began to acquire the second strategy which Anzai and Simon called the goal-peg strategy. This strategy was characterized by the selection of intermediate goals during problem solving. For example, at one point the subject indicated that one of her goals was to move disc 4 to peg B. Shortly after she evolved the goal-peg strategy, the subject developed a third strategy, called the recursive subgoal strategy. This strategy involved recognizing relationships among various newly established goals that the subject had constructed. The subject
seemed to be reasoning that in order to achieve goal A it was necessary to master goal C, and so forth. The final strategy employed by the subject was termed the pyramid subgoal strategy. In developing this strategy the subject learned to view problem goals in a new way. Goals in this strategy involved transferring discs in sets rather than one at a time. The name pyramid presumably derived from the fact that any given set of discs placed on a peg took the shape of a pyramid.

After analysis of the subject's protocol, Anzai and Simon developed a computer program to simulate each of the identified strategies. The initial simulation involved a fixed production system. No attempt was made to model the learning process at this stage. The fixed system afforded a way to compare productions associated with the various strategies employed by the subject. These comparisons made it clear that in adopting each new strategy the subject used information acquired in the course of the problem solving with the previous strategy. For example, the selective search strategy revealed the necessary steps to achieve problem-solution and made it possible to eliminate unnecessary steps. The ability to focus on necessary steps provided the information used in identifying subgoals leading to overall problem solution. The identification of subgoals in turn made it possible to conceptualize relationships among subgoals.
Following simulation using the fixed production system, Anzai and Simon developed an adaptive production system to simulate the subject's learning. This system incorporated the following assumptions. First, it presumed that fundamental to learning is the ability of the system to acquire knowledge and to use that knowledge to modify itself. Second, having recognized that a particular action outcome is either bad or good, the system is capable of creating new production to modify its behavior in a desirable direction. To do this, it reasons backward from the outcome in question to find a pattern that can be hypothesized to have been responsible for the outcome. Using these assumptions, Anzai and Simon were able to create a system that employed information acquired from execution of primitive strategies such as selective search to develop higher order strategies such as the pyramid strategy. In this way the Anzai and Simon system learned in the course of problem solving.

Recently J. H. Larkin (1979) has developed an adaptive production system called ABLE that learns in the course of solving physics problems. When initially presented with a problem a version of the system called barely ABLE produces a solution using generalized methods that require considerable search. However, in the process of solving the problem the program stores specific methods that assist in working other similar problems, thus it becomes what Larkin calls a more ABLE system.
The ABLE system includes permanent knowledge encoded as productions involving condition/action pairs. The condition portion of these pairs are matched against the contents of information in short-term memory. When a condition is satisfied, the model executes the associated action. Such execution typically involves the addition and/or deletion of elements from short-term memory. Consequently, the contents of short-term memory are continuously changing. As a result, the system is responsive to a changing array of circumstances.

ABLE is capable of solving a number of specific kinds of physics problems. The familiar inclined plane type of problem affords an example. In problems of this kind, the learner's task is typically to determine the speed of an object when it reaches a certain point while sliding down the inclined plane. In arriving at a solution to the problem the learner must take into account both the force of gravity operating on the object and the force of friction impeding the progress of the object down the plane.

In order to solve a problem, the information processing system must have an internal representation of the problem (Newell and Simon, 1972). ABLE represents problems in a manner analogous to the rough sketches learners sometimes draw as they initially meet a problem. For example, in solving the inclined plane problem a learner might draw a sketch including a plane and an object located on the
plane. The sketch would include information about the position of the object, its motion, and its speed.

In categorizing information in problem representation, ABLE labels the variables with which it is working as either known or desired. For instance, the position of the object on the inclined plane may be known while the speed is desired.

Problem solving is initiated by implementing the barely ABLE version of the system. The barely ABLE model uses a very primitive selection strategy. It notes that a given quantity is desired and searches its memory for a principle that includes the quantity. After relating each variable in the equation on which it is working to a known or desired quantity, it notes what quantities have not yet been accounted for in the problem. These are listed as desired and the system then searches for new principles that may include them. When all the variables within an equation except one are known, the system concludes that it can solve the equation directly for the remaining unknown.

Although barely ABLE is not a very effective problem solver, it does include one very important redeeming feature. When it finds an equation it can actually solve, it makes a record of the circumstances and the result. It stores both the known and the desired quantities involved in arriving at a solution. Consequently, when and if the model encounters the circumstances that lead to the execution of
problem solving involving the equation in question on a subsequent occasion, it is able to achieve a solution in a single step simply by recalling information stored in long term memory. ABLE's learning then involves building up through experience large quantities of highly specific knowledge useful in particular situations. Larkin refers to the knowledge that ABLE acquires as automatic knowledge because this type of knowledge is able to generate new information automatically from available knowledge, without the application of any specific problem-solving strategy.

As barely ABLE is given the opportunity to practice on many problems it becomes more ABLE. Its change in competency results first of all from a reduction in the amount of search activity necessary in problem-solving. As a result of learning ABLE acquires detailed patterns for recognizing situations for which a given principle may be useful. Not only does practice make it possible for ABLE to reduce search but also the new productions developed during the course of learning are collapsed or compiled entities involving several productions from the barely ABLE version of the model.

Larkin assessed the extent to which her model corresponded to the behavior of human problem solvers by using a protocol approach similar to that employed by Anzai and Simon (1979). Larkin asked both novice and expert problem solvers to work physics problems such as the inclined plane
problem. The learners were instructed to think aloud during the course of problem solving. The novice problem solvers were beginning students at the University of California at Berkeley and the experts were professors. Larkin found that novice problem solvers exhibited behavior similar to that of the barely ABLE version. However, the novices did incorporate some of the strategies of the more advanced system. She suggests that the behavior of novices can be modeled by incorporating variants of the barely ABLE system which she calls slightly ABLE versions of the model. The more ABLE version gave a rather close correspondence to the performance of the expert subjects. However, there were slight differences in the solution paths followed by the experts and the more ABLE version of the system.

J.R. Anderson (1980) has developed an adaptive production system called ACT which he has used to construct a general theory of learning. Anderson has applied his system to the study of problem-solving in geometry. In particular he has applied the system in problems involving proof construction. For example, the system can solve problems of the type in which a learner is asked to prove that two triangles are congruent on the basis of information about some of the sides and angles included in the triangle.

Anderson assumes that proof generation in geometry involves two kinds of activities. In the first of these, the learner makes a plan for executing the proof. In the
second, the proof is actually carried out. These two kinds of activities are not necessarily executed consecutively. Rather, there is typically some alteration back and forth between the two.

Planning activity typically involves the construction of an outline that includes specification of a set of geometric rules that allow the learner to progress from the problem givens through intermediate statements to the to-be-proven statement. Anderson uses the term proof-tree to refer to plans evolved as a result of planning activity. Finding a proof-tree is accomplished by searching forward from the givens to the to-be-proven statement and/or by searching backward from the to-be-proven statement to the givens. ACT executes both forward and backward searches simultaneously. Anderson observed that human problem solvers also appear to employ both forward and backward search in generating a proof-tree.

Learners typically receive the information used in proof construction in what Anderson called declarative form. For example, definitions, postulates, and theorems are usually presented as verbal declarative statements. In order to apply information of this kind it is necessary for the learner to relate it to the specific situation or set of conditions involved in the problem-solving endeavor. In the ACT system, when a particular piece of declarative information is used in problem solving, a new production is formed
that enables the system to apply the information acquired on future occasions. The mechanism of converting declarative information into productions that can be used to solve specific problems is called procedural compilation.

Procedural compilation in ACT involves two major processes: composition and instantiation. When a series of productions occur in a fixed order in proof construction, composition creates a new production that achieves the goal of the initial sequence in a single step. Note the similarity between this process and techniques used by Larkin in her ABLE production system. Instantiation is a process that constructs specialized versions of productions in which information that would otherwise have to be retrieved from long-term memory is included as an integral part of the production. Instantiation makes it possible to apply larger productions by eliminating the amount of information that must be retrieved from long-term memory.

The knowledge compilation processes of composition and instantiation are compatible with a number of phenomenon observed in human learning. For example, in the process of making a skill automatic through practice it is frequently observed that there is some loss of awareness in executing the skill. This is congruent with the fact that given the application of compilation processes, knowledge is no longer being retrieved into short-term memory as it would otherwise have to be.
Adaptive Production Constructs and Fraction Identification—A fundamental feature of the adaptive production approach to the explanation of learning is the idea that information processing systems can learn from their own actions. Each of the production systems described above uses the strategy of recording various conditions and actions producing desired outcomes in problem solving. This stored information is then used to form new productions. In every case the fundamental assumption underlying construction of a new production is the idea that problem solving would be made more efficient as a result of the addition of the new production. Typically gains in efficiency involve some new way to eliminate activities that would otherwise have to be performed by the system.

It is assumed that learners acquiring skills in fraction identification under the type of training planned in the present study will display learning processes similar in some ways to those illustrated in the adaptive production system described above. Nonetheless, there are important differences between the adaptive production examples cited above and fraction identification learning as conceived in the present investigation. One of these differences relates to the fact that the fraction identification tasks to be examined in the present study are algorithmic in nature. As already indicated, the search requirements associated with solving fraction identification problems are minimal.
The problems can be solved by executing a uniform set of actions in a relatively fixed sequence. By contrast, the problem-solving tasks examined in adaptive research have placed extensive search demands on the information processing system.

A second difference between the adaptive production systems discussed above and the current fraction identification learning situation is that in the present study correct problem solving is modeled for the learner whereas the adaptive production examples required the learner to discover the solution to the problem to be solved. Much has been made in the educational literature of the distinction between discovery learning and learning under demonstration. Learning under demonstration is often described as passive learning in which the learner simply receives information from an external source rather than discovering something for himself or herself.

Despite the differences mentioned here it may be that certain of the kinds of learning mechanisms illustrated in adaptive production research will operate in the current fraction identification case. More specifically, it is reasonable to assume that learners will attempt to refine their fraction identification efforts by eliminating certain steps in the identification process. For instance, learners confronted with the one-element exemplars during training may come to realize that set division is not essential to
correct problem solution even though the process of set division is repeatedly demonstrated by the experimenter. If a child is asked to identify five-sixths of six objects, she/he may skip the step of dividing the six element set into six equal parts. Under these conditions the child would mark five objects of the set because the number in the numerator was five. A response of this kind would support the acquisition of a fraction-identification approach congruent with the one-element rule. As indicated previously, learners operating in accordance with that rule solve fraction identification problems by putting the letter X on the number of the objects designated by the fraction numerator.

Learning occurring under denominator-rule exemplars could also involve the elimination of set division. A child might gain some efficiency in fraction identification by learning that correct solutions might be achieved by equating the numerator of the fraction to be identified with the number of columns in the set of objects presented. For example, to identify two-fifths of ten objects the child might count two columns and place a letter X on all objects in those columns. This strategy would yield results consistent with the denominator rule, but nonetheless eliminate set partitioning.

As indicated in Chapter 1, the present study will include a separate investigation that will provide an opportunity to observe the kinds of problem-solving procedures
which the children actually adopt in solving fraction identification problems. This phase of the study is intended to afford increased insight into the manner in which children acquire competencies consistent with the one-element and denominator rules.

Information Processing Theory and Methods of Research

As illustrated in the preceding section, the methods used by information processing theorists in the conduct of research are quite different from those generally employed in experimental studies in the field of psychology. Information processing theorists generally study the behavior of a small number of subjects intensively. They typically collect protocols of the subject's thinking as revealed by verbal reports taken during the course of the cognitive activity. The reports are often related to computer simulation of cognitive functioning.

The methods adopted in information processing research are associated with the manner in which information processing researchers conceptualized psychological theory and the scientific methods used to provide the empirical evaluation of the theory. This section will review information processing concepts of psychological theory and the basis for the methods they employ in studying psychological phenomena.
The Nature of Theory from an Information Processing Perspective—Information processing theorists use different criteria for evaluating theory than those traditionally employed in experimental research. In order for a theory to be plausible from an information processing perspective it must include three conditions (Miller, 1978). First, it must be possible to show that the mechanisms assumed to operate within the theory could actually occur from a physical standpoint. For instance, it would be futile to construct a theory of human information processing that required the memory capacity and speed of processing ordinarily associated with a high speed digital computer. The second condition is that it must be feasible to state the rules that govern interactions among the mechanisms proposed in the theory. Finally, it must be possible to show that the result of the interactions eventuates in the performance of the cognitive function that is being explained. There are many possible ways in which the three conditions mentioned here may be met. The most frequently used method is to construct computer programs which conform to these conditions.

Adoption of the three conditions presented above represents a commitment to sufficiency in theory construction. Information processing theorists hold that there are few if any adequate psychological theories currently available for explaining cognitive functioning. The reason is that they
believe that existing theories are not sufficient to explain behavior. For instance, to say that behavior is a function of reinforcement is not an adequate psychological theory from the information processing perspective because that statement affords insufficient information to make the behavior in question occur. This is easily demonstrated by the fact that in order to simulate even the simplest form of cognitive functioning a computer program would have to incorporate many principles beyond the reinforcement principle stated in the theory.

Because information processing theorists take the view that there is currently a scarcity of psychological theories adequate to explain cognitive functioning, they place great stress on theory development (Miller, 1978). Little emphasis is given to the time honored task of illustrating that a given theory is useful. Accordingly, hypothesis formation and associated hypothesis testing as they almost invariably occur in experimental psychology are not typically employed in information processing studies.

There is no necessary conflict between the hypothesis testing approach that has been the mainstay of experimental research in the past and the theory development stance employed in information processing studies. They simply serve different purposes. The present investigation will utilize both kinds of procedures. The main portion of the study will incorporate hypothesis testing procedures of
the type typically used in experimental investigations. However, in addition, an exploratory phase will be included in the investigation that will involve the collection and intensive study of verbal protocols of some of the children engaged in the act of learning fraction identification. It is hoped that the verbal protocols will provide additional insight into the results obtained through hypothesis-testing procedures. Moreover, the information gained from the protocols may form the basis for future research aimed at theory development of the type engaged in by information processing theorists.

Verbal Reports and Information Processing Research—As illustrated in the research described above, information processing theorists rely heavily on verbal reports to provide evidence regarding knowledge structures. With the advent of the behaviorist position during the early part of the twentieth century (Watson, 1913), verbal reports became suspect as data in the conduct of psychological research. Suspicions about verbal data have remained until the present time (Nisbett and Wilson, 1977). The prevalent view in many circles is that verbal reports provide potentially interesting but only informal information which is worthless for hypothesis verification (Ericsson and Simon, 1980). This view has tended to foster a somewhat careless attitude about verbal report methodology. Until recently no
clear guidelines existed for distinguishing between introspective verbalizations of questionable utility and verbal output that might be regarded as useful data for providing support for hypotheses related to cognitive functioning. Moreover, no distinctions were made within the field of psychology among various forms of verbal data.

Recently Ericsson and Simon (1980) have conducted an extensive survey of the literature on verbal report data. Their analyses of verbal data provide many guidelines for distinguishing among useful verbal information and verbal data of questionable utility. Ericsson and Simon describe a number of significant dimensions along which verbal report data can vary. They then relate these dimensions to the possibility that verbal information may distort accuracy in reporting cognitive functioning.

One of the major dimensions identified by Ericsson and Simon (1980) involves current versus retrospective verbal information. Information regarding cognitive activities verbalized at the time of occurrence of the cognitive process under study is described as current data. For example, if a learner solving a fraction identification problem were to say "Now I am going to divide the set into two parts" that would be regarded as a verbal description of currently attended to cognitive activity. By contrast, information about cognitive functioning that has occurred previously is described as retrospective. For instance, in response to a
question such as "What are you thinking about when you solved the last problem?", a child might say "I wanted to make sure that I divided the sets into parts."

Ericsson and Simon (1980) cite a number of studies indicating that when the current verbalizations are used to describe the information currently available in short-term memory, there will be no distortion in the course or structure of the cognitive process under investigation. By contrast, information reported in retrospect may be distorted because of failure to recall the cognitive activity under study accurately.

A second dimension along which verbal report data might vary is related to the nature of the report (Ericsson and Simon, 1980). The extremes of this dimension are reports about general cognitive functioning versus reports about particular occurrences in cognitive activity. This dimension is of particular importance with respect to retrospective verbal reports. Generalized reporting is frequently used because subjects may not remember accurately specific processes after a significant amount of time has elapsed following the occurrence of the cognitive activity under study. Ericsson and Simon point out that questioning regarding general thought processes necessarily involves interpretations on the part of the subjects. Interpretative responses require inference that often cannot be relied upon to provide accurate information about cognitive functioning.
A third dimension along which verbal reports may vary involves the nature of the probes which are employed to stimulate verbalization (Ericsson and Simon, 1980). Ericsson and Simon describe some probes as directed and others as undirected. The question "Did you divide the set into parts?" is an example of a directed probe. "What are you thinking now?" would be an example of an undirected probe. Ericsson and Simon point out that directed probes may predispose the subject to respond in particular ways and thereby distort the information about cognitive functioning.

Ericsson and Simon's (1980) review of the literature on verbal report data supports the use of the so-called think-aloud technique in examining cognitive processes. This method requests the subject to think aloud as he or she engages in cognitive activity. The procedure has the advantage of relying only on information currently in short-term memory. Thus, inaccuracies that might result as material is retrieved from long-term storage are avoided. Moreover, the need for retrospective inferences that might lead to distortion in data about cognitive functioning is circumvented. For these reasons, the present research will make use of the think-aloud procedure in the exploratory stage of the investigation.
CHAPTER 3

METHOD

Subjects

Eighty-two subjects from elementary schools in Tucson, Arizona participated in this investigation. Subjects were enrolled in grades one through four and ranged in age from six to ten years. The children came from middle and lower class areas of the community and were predominantly from the Mexican-American and Anglo populations. Table 1 details age, grade, sex and ethnicity information for the sample.

Table 1. Age, Grade, Sex and Ethnicity of Study Participants

<table>
<thead>
<tr>
<th>Age in Years</th>
<th>Number of Cases</th>
<th>Grade</th>
<th>Number of Cases</th>
<th>Sex</th>
<th>Number of Cases</th>
<th>Ethnicity</th>
<th>Number of Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>17</td>
<td>1</td>
<td>22</td>
<td>Male</td>
<td>38</td>
<td>Anglo</td>
<td>62</td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>2</td>
<td>10</td>
<td>Female</td>
<td>44</td>
<td>Mexican-American</td>
<td>18</td>
</tr>
<tr>
<td>8</td>
<td>21</td>
<td>3</td>
<td>28</td>
<td></td>
<td></td>
<td>Other</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>25</td>
<td>4</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Experimenters

The three experimenters involved in the study were graduate students enrolled at The University of Arizona. They served on a volunteer basis and received remuneration for their services. Each experimenter underwent approximately two hours of simulated training at The University of Arizona. The simulated training involved the examination of a complete set of training and recording materials, verbal explanation of the training and recording procedures, and participation in practice training with other experimenters and with young children. Experimenters were assigned randomly to school settings, trained approximately the same number of children and trained children randomly assigned to each of the experimental conditions. Experimenters were observed periodically during the course of the study to ensure adherence to specified procedures.

Materials

The materials used in this study consisted of several pencil and paper tasks, each of which was printed on standard sized paper and was contained within booklets. Included were a pretest, a prerequisite skill test, training sheets and posttests. An example of each of the task materials is presented in Appendix A.
Pretest

The pretest consisted of a set of four problems representative of the problems to be included in the posttest. Students were asked to mark the letter X on the number of units which represented one-third and two-thirds of sets containing both three and six elements. These fractions were chosen as test fractions because they served to screen children with prior knowledge of the task to be taught. In addition, the pretest fractions were small fractions which could not be reduced to lower terms. All elements within a set were one of three shapes: triangle, circle, square. Each shape appeared in approximately the same number of problems across the pretest.

Prerequisite Skill Test

The prerequisite skill test consisted of eight problems related to subset division. Students were asked to divide sets of nine units into three equal subsets, sets of 12 units into four equal subsets, sets of 15 units into five equal subsets, and sets of 18 units into six equal subsets. These set and subset sizes were chosen because they include the sets and subsets related to multiple-element problems appearing in the project tasks. Set units consisted of one of three shapes (triangle, circle, square); shapes were used with approximately equal frequency.
Training Sheets

There were eight training sheets for each training group. Each sheet consisted of a single problem related to one of the four training fractions and appropriate to the exemplar type for the specific training group. Training fractions included one-fourth, three-fourths, one-sixth and five-sixths. Training fractions were chosen because they differed from the fractions included in testing materials and because they were both small in size and could not be reduced to lower terms. Set units consisted of one of three shapes (triangle, circle, square); each shape appeared with approximately equal frequency across training sheets.

Posttests

Two posttests were used in this investigation. Each of the posttests contained eight items which were consistent in number with one of the denominator types, that is, consistent with either single-unit or multiple-unit subsets. Each problem within the posttests consisted of a request to mark with the letter X a specific fractional part of the units in a pictured set. Test fractions included one-third, two-thirds, one-fifth and two-fifths. These fractions were chosen because they were small fractions and because they could not be reduced to lower terms. The two posttests were presented as a single unit in which problems representing the two
denominator types were randomly ordered. Set units, as in previously discussed materials, were one of three shapes.

Procedures

Pretest

A pretest designed to measure fraction identification capability was administered in a group setting to each child to determine his/her eligibility to participate in the investigation. Children answering none of the pretest problems correctly were considered eligible participants and were then tested individually for prerequisite skill capability.

Prerequisite Group and Training Assignment

Participants' prerequisite skill scores were used to assign them to prerequisite skill groups. All children correctly answering fewer than four of the eight items on the prerequisite test were assigned to the first prerequisite group and then were randomly assigned to one of the two training groups while children correctly answering four or more problems on the prerequisite test were assigned to the second prerequisite group and then randomly assigned to the training groups.

Training Group Variations

The two experimental training groups were involved in a fraction identification program consisting of eight separate training trials. Each trial was based on one of
the following fractions: one-fourth, three-fourths, one-sixth and five-sixths. Fractions were presented to students individually and in random order.

Learners were presented training consisting of both symbolic and behavioral modeling and of evaluative feedback. In this study symbolic modeling consisted of verbal statements of the rules pertinent to the task at hand. A fraction identification problem involves knowledge of three rules: (1) definition of the fraction; (2) partitioning of the set into nths where n is the size of the denominator; and (3) marking the numerator-specified subsets. The experimenter verbal statements related to these rules were of the following format: "Three-fourths means three parts of four parts. First we need to make four equal parts. One. Two. Three. Four. Now we need to mark the shape(s) in one part with the letter X. One part. Two parts. Three parts."

Experimenter behavioral modeling consisted of behavioral demonstration of the rule statements. For example, following symbolic modeling of the partitioning rule, the experimenter partitioned an exemplar by drawing vertical lines between subsets. Following the statement of the numerator rule, the experimenter marked the appropriate number of set units with the letter X.

After each instance of training, the learner was presented with a fraction-identification task identical to the one just modeled by the experimenter. The learner's
performance on this task received one of two evaluative feedback statements: (1) "Yes. That is exactly right. Now let's do another.", or (2) "That is not quite right. Let's try another." A new trial was presented immediately following the evaluative feedback.

Although the basic format was exactly the same for the two experimental groups, the groups varied on the nature of the exemplars which they received during training. Group 1 received exemplars which were ambiguous in nature. That is, the exemplar sets provided Group 1 learners allowed correct fraction identification through the use of either the stated denominator rule or the unstated one-element rule. Group 2 received unambiguous exemplars. The exemplar sets for this group were constructed so as to demand the use of the denominator rule during fraction identification. Inaccurate identification would occur if the one-element rule were used with Group 2 exemplar sets.

Posttest Problem Type Variations

Subsequent to training, each participant was asked to respond to 16 fraction identification problems, eight of which could be correctly marked by applying either the one-element or the denominator rule and eight of which required use of the denominator rule for accurate identification. As discussed earlier, questions related to the two problem types were encompassed within a single booklet and were arranged in
random order within the booklet. As can be noted from the posttest example in Appendix A, the one-element problems were item numbers 1, 3, 6, 11, 13, 14, 15; the remaining items were denominator-rule problems.

Exploratory Phase

In an effort to gain both additional information to corroborate the assumptions examined through formal hypothesis testing and further insight into the manner in which children solve fraction-identification problems, two children were selected to participate in the exploratory phase of this investigation. Children participating in this phase of the investigation received training identical to that provided the children in the main part of the study. These children were asked to think aloud as they were trained and as they took the posttests. Indirect probes were employed to elicit additional information. Tape recordings were made of the learners' verbalizations.
CHAPTER 4

RESULTS

Two hypotheses set forth in this investigation were related to the association between the prerequisite skill of set partitioning and rule acquisition. Examination of the data revealed that very few individuals (less than 10% of the students eligible to participate in the study) had mastered the hypothesized prerequisite skill. Consequently, it was decided to forego statistical analysis of this variable. Discussion related to set partitioning as a prerequisite to fraction identification is contained in Chapter 5.

A review of the children's responses to both of the posttests measuring fraction identification skills suggested that for both tests student performance reflected an almost perfect dichotomy. On Posttest 1, which contained problems related to the one-element rule, 79 of 82 students (96%) obtained scores of zero, seven, or eight. On Posttest 2, in which correct fraction identification required use of the denominator rule, 75 of the 82 students (92%) answered none, seven, or eight of the items appropriately. Because fraction identification performance clearly did not reveal an interval scale, it was deemed appropriate to adopt categorical data analysis procedures to analyze the results.
To employ categorical data analysis procedures, each of the two types of fraction identification tasks was dichotomized. A score of one was assigned for children who attained seven or more correct responses. A score of two was given for less than seven appropriate identifications. Of the 53 failing scores across the two posttests, 42 (80%) were given for zero correct responses while only two (3.7%) were given for 50% right.

**Modified Path Analysis Assessing Training Effects**

A modified path analysis procedure developed by Goodman (1973) was used to assess the effects of training group assignment on fraction identification performance. The Goodman procedure involves the application of logit models in contingency table analysis. The contingency table under examination in the present study was comprised of the cross-classification of three dichotomous variables. The first represented training group assignment. Individuals exposed to one-element exemplars were assigned to Group 1, while individuals exposed to the denominator-rule exemplars were assigned to Group 2. The second variable reflected pass/fail performance on one-element problems and the third represented pass/fail performance on denominator-rule problems. For both of these variables a passing score was coded one and a failing score was coded two.
Goodman (1973) has shown that it is possible to analyze the effects of one or more categorical variables on a dichotomous dependent measure through the use of logit models similar to analysis of variance models. In these models, the dependent measures are treated as logits. A logit is defined as the natural logarithm of the odds that a given dependent measure will occur at one level as opposed to the other. For example, a logit for the denominator-rule posttest could be constructed by dividing the number of individuals estimated under a given model to respond correctly to denominator-type problems by the number of individuals expected under the model to respond incorrectly. The natural logarithm of this ratio would be the logit for denominator-rule problems. The logarithm is used to make the model additive and thus similar to the analysis of variance. However, meaningful interpretations of results can be made without the logarithmic transformation. When the logarithmic transformation is omitted, a multiplicative rather than an additive model results.

Logit models to test hypotheses related to the data in a given contingency table can be hierarchically ordered. In the case of the three-way table, three kinds of models are of interest. If the variables are labeled A, B and C with C as a logit variable, the first hypothesis of interest is that of marginal independence. It asserts that the joint variable AB is independent of variable C. This hypothesis
asserts that neither A nor B affects C. The second type of hypothesis of concern asserts conditional independence among the variables of interest. One such hypothesis would assert that variable A would be independent of variable C within categories of B. In this hypothesis, the relationship between A and C is assumed to be conditional on the level of B. The hypothesis asserts that when the relationships between A and B and between B and C are taken into account, A will have no effect on C. An analogous hypothesis could, of course, be constructed for the effects of B on C. The final hypothesis of interest in the three-way table is the hypothesis of no-three-way interaction. With C as a logit variable, this hypothesis asserts that the level of association between A and B when C is at level one will equal the level of association between A and B when C is at level two. This hypothesis assumes that both A and B have a direct effect on C, but that A and B do not interact in affecting C. The hypothesis of no-three-way interaction is analogous to the hypothesis of no AB interaction in analysis of variance when variable C is an interval scale dependent measure.

Expected cell frequencies for the hypotheses related to marginal independence, conditional independence and no-three-way interaction are built by fitting the marginals of the contingency table. To fit the marginals means to make the expected frequencies equal to the observed frequencies for specified marginals. For example, Figure 1 presents a
contingency table for the hypothesis of marginal independence. In this model, the variables A and B are treated as a joint variable and C is alone. The hypothesis which is being tested when AB is a joint variable relates to the independence or lack of independence existing between this joint AB variable and the C variable. The nature of the relationship between variables A and B is not addressed in this logit model. Logit models designed to evaluate hypotheses of conditional independence are built by fitting with two joint variables. In the example shown in Figure 2, which depicts the contingency tables appropriate to test the independence between variables B and C within levels of A, the fitted marginals involve the joint variables AB and AC. To test the no-three-way interaction hypothesis, the marginals for joint variables AB, AC and BC are fitted. The no-three-way interaction hypothesis may be thought of in terms of tables similar to those in Figure 2. Under this hypothesis it is assumed both that the level of association between each of two tables will be equal and that equality between pairs of tables will hold across variables. That is, the A C relationship will be the same within levels of B and the B C relationship will be the same within levels of A.

Logit models may be tested using the chi-square statistic to assess the correspondence between observed cell frequencies and estimates of expected cell frequencies generated under the model. Fay and Goodman (1973) have
Figure 1. Marginal Independence Contingency Table - Hypothesis: Joint Variable AB Independent of Variable C

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>Fitted Margins</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>F_{111}</td>
<td>F_{112}</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>F_{211}</td>
<td>F_{212}</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>F_{121}</td>
<td>F_{122}</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>F_{221}</td>
<td>F_{222}</td>
</tr>
</tbody>
</table>

F_{..1} F_{..2}

Figure 2. Conditional Independence Contingency Table - Hypothesis: Variable B is Independent of Variable C within Levels of A

<table>
<thead>
<tr>
<th>A_1</th>
<th>A_2</th>
<th>Fitted Margins</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>C</td>
<td>AC</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>F_{111}</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>F_{221}</td>
</tr>
</tbody>
</table>

F_{1.1} F_{1.2} F_{2.1} F_{2.2}
developed a computer program called ECTA (Everyman's Contingency Table Analyzer) that can be used to generate expected cell frequencies for logit models and that carries out appropriate chi-square tests. The ECTA program was used in the present study.

As indicated, the logit models described above are hierarchically ordered. Two models are hierarchically ordered when one contains all of the constraints of the other plus one or more than one additional constraints (Goodman, 1973). For example, the hypothesis that A is independent of C within categories of B contains all of the constraints imposed under the hypothesis that the joint variable AB is independent of C and an additional constraint. This additional constraint eventuates in the loss of one degree of freedom in model testing.

Hierarchical models can be compared statistically (Goodman, 1972). The likelihood ratio statistic is particularly useful in model comparison because it partitions exactly (Bishop, Fienberg and Holland, 1975). To compare two models the $X^2$ for the one with the smaller number of degrees of freedom is subtracted from the $X^2$ for the one with the larger number of degrees of freedom. The result will be a $X^2$ with degrees of freedom equaling the difference between degrees of freedom for the two models being compared. The resulting $X^2$ can be referred to a table for the chi-square
distribution to test the hypothesis that one model improves significantly on the fit afforded by the other.

Table 2 shows the results of the Chi-Square tests for the modified path analysis conducted to assess training group effects in the present study. Eight hierarchichal models were tested in the analysis. These are designated as models $H_0$ through $H_7$ in the table. The $X^2$ values, degrees of freedom and p values are indicated for each of the models tested.

**Table 2. $X^2$ Values for Eight Models of Independence**

<table>
<thead>
<tr>
<th>Model</th>
<th>Fitted Marginals</th>
<th>$X^2$</th>
<th>Degrees of Freedom</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0$</td>
<td>A, B, C</td>
<td>32.79</td>
<td>4</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>$H_1$</td>
<td>AB, C</td>
<td>30.48</td>
<td>3</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>$H_2$</td>
<td>AC, B</td>
<td>7.38</td>
<td>3</td>
<td>.060</td>
</tr>
<tr>
<td>$H_3$</td>
<td>BC, A</td>
<td>31.43</td>
<td>3</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>$H_4$</td>
<td>AB, AC</td>
<td>5.08</td>
<td>2</td>
<td>.079</td>
</tr>
<tr>
<td>$H_5$</td>
<td>AB, BC</td>
<td>29.12</td>
<td>2</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>$H_6$</td>
<td>AC, BC</td>
<td>6.03</td>
<td>2</td>
<td>.049</td>
</tr>
<tr>
<td>$H_7$</td>
<td>AB, AC, BC</td>
<td>1.12</td>
<td>1</td>
<td>.500</td>
</tr>
</tbody>
</table>

Variable A in the table is training group; variables B and C are the two fraction identification posttests. Variable B is related to fraction identification demanding only the one-element rule while variable C is related to problems requiring use of the denominator rule.
The aim of the chi-square analysis is to find a preferred model which both affords an acceptable fit for the data and a parsimonious explanation of the findings. High $X^2$ values indicate a poor fit of the model to the data. On the other hand, low $X^2$ values indicate a close correspondence between observed and expected cell frequencies and hence a good fit for the data. A parsimonious explanation refers to an explanation containing as few variables as possible while obtaining a good fit.

Model $H_0$, the mutual independence model, asserts mutual independence among the three variables under examination. The $X^2$ of 32.79 with 4 degrees of freedom and a $p$ value of less than .001 indicates that this model fits the data very poorly. Models $H_1$ through $H_3$ reflect the marginal independence hypotheses depicted in Figure 1. Model $H_1$ asserts that the joint variable $AB$ (training group, one-element posttest) is independent of variable $C$ (denominator-rule posttest). The $X^2$ of 30.48 with 3 degrees of freedom and a $p$ value of less than .001 indicates that this model also fits the data very poorly. Models $H_2$ and $H_3$ are similar to $H_1$ in that each asserts independence between a joint variable and third variable. Model $H_2$ asserting independence between the joint variable $AC$ (training group, denominator-rule posttest) and variable $B$ (one-element posttest) affords a marginally acceptable fit for the data. The $X^2$ value of
7.38 with three degrees of freedom obtained for model $H_2$ has a $p$ value of .06.

Models $H_4$ through $H_6$ hypothesize conditional independence with respect to the variables in the study. Figure 2 depicts the contingency tables relevant to the conditional independence hypothesis. Model $H_4$ asserts that the performance on one-element items is independent of performance on denominator-rule items within each of the two training group categories. The $X^2$ of 5.08 with two degrees of freedom obtained for model $H_4$ has a $p$ value of .08 which suggests that this model affords an acceptable fit for the data. Models $H_5$ and $H_6$ are similar to model $H_4$. Model $H_5$ asserts that training group assignment is independent of performance on the denominator-rule problems within categories of performance on the one-element problems. With a $X^2$ of 29.12, two degrees of freedom and a $p$ value of less than .001, this model provides an inadequate fit with the data. Model $H_6$ asserts that performance on one-element problems is independent of training group assignment within the two levels of performance within denominator-rule problems. The $X^2$ of 6.03 with two degrees of freedom and a $p$ value of less than .05 indicates that this model is marginal with respect to the fit it provides for the data.

Model $H_7$ is the model of no-three-way interaction. This model reflects effects of training group on fraction identification and association between the two types of
fraction identification. This $X^2$ of .12 with one degree of freedom and a $p$ value of greater than .5 indicates that this model fits the data very well.

As the above results show, more than one model provides an acceptable fit for the data in this study. Because more than one model provides an adequate fit for the data it is necessary to determine a preferred model from among them. A preferred model is one that affords a significantly better fit for the data than that obtained with other models containing fewer variables. To determine which model is the preferred model, hierarchical comparisons must be made. As described above, hierarchical comparisons are made by subtracting the $X^2$ for a model which fits the data from that of another hierarchically related model which also fits the data and which contains more degrees of freedom than those in the subtrahend. The significance of the resulting $X^2$ is then determined.

In the present study the model of no-three-way interaction, model $H_7$, is hierarchically related to each of the other models tested. The subtraction of the $X^2$ for model $H_7$ from any of the other $X^2$s reveals a significant improvement in the fit of the model to the data. Moreover, as indicated above, model $H_7$ provides a very good fit for the data. Because model $H_7$ significantly improves on the fit afforded by each of the other models and because it fits the data well,
it was adopted as the preferred model for explaining the association in the table.

By accepting model H, as the preferred model, the effects of training group on one-element posttest and denominator-rule posttest performance are accepted. Also accepted is the assumption that each pair of variables maintains the same relationship across levels of the remaining variable. For example, under this model the relationship between training group and the denominator-rule posttest is best described as constant across levels of the one-element rule posttest. Also under this model the relationship between training group and the one-element rule problems remains consistent across levels of the denominator-rule posttest.

Once a preferred model has been selected, analysis of the effects related to the model can be undertaken. One way to investigate the effects of training and the posttests is to express model H, as a logit model, that is, a model in which the dependent variable is treated as a logit. It will be recalled that a logit is the natural logarithm of the odds that a given dependent variable will occur at one level as opposed to another. The following equations developed by Goodman (1973) express the results of model H, as a logit model.

$$ \phi_{ij} = \beta_{i} + \beta_{j} + \beta_{ij} \quad (1) $$
In the first of these equations, variable C, representing performance on denominator-rule problems, is treated as the logit variable. The symbol, \( \phi_{ij} \), indicates the natural logarithm of the odds that variable C will be at level 1 as opposed to level 2 when variables A and B are at levels i and j, respectively, \( (i = 1, 2; j = 1, 2) \). The second equation treats variable B, representing performance on one-element problems as the logit variable. The symbol, \( \phi_{i.k} \), indicates the natural logarithm of the odds that variable B will be a level 1 as opposed to level 2 when variables A and C are at levels i and k, respectively.

Equations 1 and 2 are analogous to the type of equations found in analysis of variance or in least squares regression. For example, in equation 1, \( \beta^C \) represents the general mean for variable C, \( \beta^{AC} \) indicates the effect of variable A (training group) on variable C (denominator-rule problems) and \( \beta^{BC} \) indicates the effect of variable B (one-element problems) on variable C. In equation 2, \( \beta^B \) represents the general mean for variable B, \( \beta^{AB} \) indicates the effect of variable A on variable B and \( \beta^{CB} \) indicates the effect of variable C on variable B.

Model H7 is represented visually in the path diagram shown in Figure 3. Figure 3 indicates that training group
Figure 3. Path Diagram Describing Relations Between Variables A, B and C

has an effect both on the one-element posttest performance and on denominator-rule posttest performance. It further indicates that performance on either of the fraction identification tasks influences performance on the other. The magnitude of effects calculated under model $H_7$ are also shown on the figure. The numerical expressions of these effects are analogous to the path coefficients in path analysis. Attention to both signs and size within the figure is appropriate. Note, for instance, that the effect of training group on denominator-rule posttest performance is negative. This indicates that assignment to the training group receiving one-element exemplars has a negative or adverse impact on the acquisition of the denominator rule.

Further light can be shed on the magnitude of effects by examining selected odds and odds ratios associated with expected cell frequencies under Model $H_7$. Table 3 contains observed frequencies, expected frequencies, odds and odds ratios related to training group effects and to the
Table 3. Observed Frequencies, Expected Frequencies, Odds and Odds Ratios for Training Groups and Problem Type Under the Hypothesis of No-Three-Way Interaction

<table>
<thead>
<tr>
<th>Response Pattern Item</th>
<th>Observed Frequency</th>
<th>Expected Frequency</th>
<th>Odds</th>
<th>Odds Ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>A* B* C*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 1 -</td>
<td>39</td>
<td>39.00</td>
<td>19.50</td>
<td></td>
</tr>
<tr>
<td>1 2 -</td>
<td>2</td>
<td>2.00</td>
<td>19.50</td>
<td></td>
</tr>
<tr>
<td>2 1 -</td>
<td>35</td>
<td>35.00</td>
<td>5.83</td>
<td></td>
</tr>
<tr>
<td>2 2 -</td>
<td>6</td>
<td>6.00</td>
<td>5.83</td>
<td></td>
</tr>
<tr>
<td>1 - 1</td>
<td>7</td>
<td>7.00</td>
<td>.21</td>
<td></td>
</tr>
<tr>
<td>1 - 2</td>
<td>34</td>
<td>34.00</td>
<td>.21</td>
<td></td>
</tr>
<tr>
<td>2 - 1</td>
<td>29</td>
<td>29.00</td>
<td>2.42</td>
<td></td>
</tr>
<tr>
<td>2 - 2</td>
<td>12</td>
<td>12.00</td>
<td>2.42</td>
<td></td>
</tr>
<tr>
<td>- 1 1</td>
<td>34</td>
<td>34.00</td>
<td>17.00</td>
<td></td>
</tr>
<tr>
<td>- 2 1</td>
<td>2</td>
<td>2.00</td>
<td>17.00</td>
<td></td>
</tr>
<tr>
<td>- 1 2</td>
<td>40</td>
<td>40.00</td>
<td>6.67</td>
<td></td>
</tr>
<tr>
<td>- 2 2</td>
<td>6</td>
<td>6.00</td>
<td>6.67</td>
<td></td>
</tr>
</tbody>
</table>

*Coding System
A - Training Group: 1=One-Element Examples
2=Denominator-Rule Examples

B - One-element Posttest: 1=Passed
2=Failed

C - Denominator-rule Posttest: 1=Passed
2=Failed
association between the two posttests. Since model H seven provided a good fit for the data and was adopted as the preferred model it is appropriate to use the expected cell frequencies to compute odds and odds ratios representing effects among the variables.

It can be seen from Table 3 that the most dramatic effects of training are related to performance on the denominator-rule posttest. The data suggest that the odds of passing the denominator-rule posttest are 11.52 times greater for children taught with denominator-rule exemplars as opposed to children taught with the simple but ambiguous one-element exemplars. Performance on the one-element posttest was somewhat poorer for children trained with denominator-rule examples than for those trained with one-element examples. Fifteen percent of the children trained with the denominator-rule examples failed to pass the one-element posttest while 5% of the one-element rule trainees failed the one-element posttest. These figures represent an odds ratio of 3.34. It can also be noted from Table 3 that for children who pass the denominator-rule posttest the odds are 17 to one of passing the one-element posttest and that the odds of passing the one-element posttest are 2.55 times greater when the denominator-rule posttest has been passed than when it has been failed.

In summary, the first stage of data analysis consisted of employing Goodman's (1973) modified path analysis
procedure to study the effects of training group assignment on fraction identification performance. The findings from this analysis are that the model of no-three-way interaction fits the data very well. That is, the analysis suggests that the association between training group assignment and performance on denominator-rule problems remains the same across both the pass and the fail levels of performance on one-element problems. This finding is analogous to a finding of no AB interaction in analysis of variance. The findings further suggest both that association between training group and performance on one-element problems remains the same across both levels of performance on denominator-rule problems and that performance on either of the fraction identification tasks influences performance on the other.

**Error Analysis**

In addition to being employed in data analysis related to training group effects on fraction identification performance, Goodman's (1973) modified path analysis procedure was also used in error analysis related to denominator-rule fraction identification problems. The hypothesis of mutual independence for the variables of training group and error type was evaluated for each of the posttest fractions. Posttest fractions included one-third, one-fifth, two-thirds and two-fifths. Error types were dichotomized as errors congruent with inappropriate use of the one-element rule
(coded one in the analysis) and as errors incongruent with
the one-element rule (coded two in the analysis). As in the
previous analyses, individuals exposed to one-element ex­
emplars were assigned to Group 1 and individuals exposed to
denominator-rule exemplars were assigned to Group 2.

Table 4 shows the results of the Chi-Square tests for
the modified path analysis conducted to assess training group
effects on error type. Four separate analyses of the mutual
independence hypothesis were conducted, one each for the
fractions one-third, one-fifth, two-thirds and two-fifths.
The obtained chi-square values indicated that in each in­
stance error type could not be considered independent of
training group assignment. That is to say, a direct rela­
tionship was found between being assigned to the training
group employing one-element exemplars and making errors con­
gruent with inappropriate use of the one-element rule on
denominator-rule fraction identification.

Table 5 contains the obtained cell frequencies for
the error data plus the odds and the odds ratios for this
data. The odds ratios have been based on observed frequen­
cies because the model of mutual independence, the only
independence model that can be investigated when looking at
the relationship between just two variables, was not found
to fit the data. As can be seen in Table 5, in all cases for
the one-element training group, the odds are high for errors
congruent with inappropriate use of the one-element rule.
Table 4. Error Analysis for Fractions One-Third, One-Fifth, Two-Thirds and Two-Fifths. Hypothesis Evaluated: Mutual Independence of Training Group and Error Type

<table>
<thead>
<tr>
<th>Fraction</th>
<th>Number of Errors</th>
<th>$X^2$</th>
<th>Degrees of Freedom</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/3</td>
<td>44</td>
<td>13.3216</td>
<td>1</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>1/5</td>
<td>43</td>
<td>22.0754</td>
<td>1</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>2/3</td>
<td>48</td>
<td>8.2932</td>
<td>1</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>2/5</td>
<td>44</td>
<td>8.2356</td>
<td>1</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>

The odds vary in magnitude from 7.25 for the fraction one-fifth to 15.00 for the fraction one-third. In addition, it can be seen in Table 5 that the odds ratios reflect a strong tendency for errors congruent with inappropriate use of the one-element rule. These odds ratios reflect the odds of making an error congruent with the one-element rule when taught with one-element exemplars as opposed to making a congruent error when taught with denominator-rule exemplars. These odds ratios range in magnitude from 8.75 for the fraction two-thirds to 21.13 for the fraction one-third. That is to say, for example considering the fraction one-third, that the odds of making a congruent error when taught with one-element exemplars are 21 times greater than the odds of making a congruent error when taught with denominator-rule exemplars.
<table>
<thead>
<tr>
<th>Response Pattern Item</th>
<th>A*</th>
<th>B*</th>
<th>Observed Frequency</th>
<th>Odds Ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction 1/3</td>
<td>1</td>
<td>1</td>
<td>30</td>
<td>15.00</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>21.13</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>7.1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Fraction 1/5</td>
<td>1</td>
<td>1</td>
<td>29</td>
<td>7.25</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>16.92</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Fraction 2/3</td>
<td>1</td>
<td>1</td>
<td>30</td>
<td>10.00</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>8.75</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>8</td>
<td>1.14</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Fraction 2/5</td>
<td>1</td>
<td>1</td>
<td>29</td>
<td>9.67</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>9.67</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

*Coding System

A - Training Group: 1=One-Element Examples
2=Denominator-Rule Examples

B - Error Type: 1=Errors congruent with one-element rule
2=Errors incongruent with one-element rule
To summarize the findings of the error analysis, the data suggest a strong relationship between training with one-element, ambiguous exemplars and performance congruent with inappropriate use of the one-element rule.

**Teacher Interview Analysis**

Some children in the investigation did not show post-test generalization to problems similar to those employed during training. That is, two children taught with one-element examples failed the one-element posttest and 12 children taught with denominator-rule examples failed the denominator-rule posttest. In an effort to isolate individual difference variables related to these performances, classroom teachers were interviewed. The variables investigated included socioeconomic status, ethnicity, age, intelligence as assessed by the child's classroom teacher and special learning styles or needs as assessed by the classroom teacher.

Information obtained from the interviews would suggest that neither socioeconomic status nor ethnicity were related to non-generalization to posttest problems resembling the training stimuli. All the non-generalizers were students from middle-class homes. The ethnic make-up of the non-generalizing group closely resembled that of the generalizers. Had the ethnicity of one child in the non-generalizer group been altered the ethnic makeup of the
total sample would have been exactly equal to that of the non-generalizer group.

Eighty-six percent of the non-generalizers were teacher rated as being of average or greater intelligence. In fact, 78% were considered to be quite bright. Thirty-six percent of the non-generalizers were teacher rated as excellent students. Most teacher comments related to special learning problems referred to non-generalizers as students who either do not pay attention to detail (21% of the group) or do not complete tasks on their own (21% of the group). Two children were described as having poor memories and these children were also described as below average in intelligence.

Whereas the majority of the individual-difference variables examined do not appear to be related to non-generalization to posttest problems similar to those employed in training, the obtained age information does suggest that this variable may be a factor in generalization. Fifty percent of the non-generalizers were six years old and another 21% were seven. All but one of these young children were trained with denominator-rule examples. They represent 56% of the young children in the denominator-rule training group.

**Analysis of Think-Aloud Data**

Think-aloud data were collected during the training of two students. As with the children in the main portion of
the study, both children were provided the denominator-rule through symbolic modeling. In addition to the symbolic modeling the children were provided behavioral modeling. One student received behavioral modeling based on one-element exemplars. For the second child the behavioral modeling was tied to denominator-rule examples. The child receiving one-element rule exemplars passed the one-element posttest but not the denominator-rule posttest, while the child taught with denominator-rule exemplars passed both posttests. These performances resembled those of children participating in the main portion of the investigation. In the main investigation 83% of the participants receiving the ambiguous, one-element exemplars did not generalize beyond one-element problems and 76% of the students taught with denominator-rule exemplars were able to pass both one-element and denominator-rule posttests. As was also the case for the larger group of children, the child taught with one-element exemplars made errors on the denominator-rule posttest problems which were in keeping with the inappropriate use of the one-element rule. The failure of the think-aloud students to learn set partitioning skills during fraction identification training also resembled the performance of the majority of children in the main study in that 93% of those initially incapable of set partitioning did not pass a second administration of the set partitioning test.
These summary facts would suggest that the think-aloud children responded in a manner resembling that of the children in the larger portion of the current investigation. It is reasonable, therefore, to believe that their think-aloud verbalizations can provide some information relative to thinking processes related to fraction identification tasks.

Protocols for the think-aloud students are provided in Appendix B. Each child's responses were consistent across all test items. Figures 4 and 5 contain examples related to the first two posttest items for the child trained with one-element exemplars and the child trained with denominator-rule exemplars, respectively. These figures illustrate the general response pattern of the children. It can be noted that the child trained with one-element exemplars responded in a fashion consistent with the one-element rule. It can also be noted that the child trained with the denominator-rule exemplars responded in a fashion consistent with the denominator rule. Interpretations of the children's verbalizations will be found in the discussion chapter.
1. Mark an X on 1/3 of the squares:

2. Mark an X on 1/5 of the circles:

Figure 4. First Two Posttest Responses Made by Think-Aloud Student Taught with One-Element Exemplars

1. Mark an X on 1/3 of the squares:

2. Mark an X on 1/5 of the circles:

Figure 5. First Two Posttest Responses Made by Think-Aloud Student Taught with Denominator-Rule Exemplars
Overall, the findings of the present investigation provided strong support for the assumptions advanced in Chapter 1. However, one unexpected finding deserves special comment. This was the finding for the prerequisite skill, subset division. Despite the fact that the present sample covered a broad age span (ages six through ten), the number of children who could perform subset division and yet could not carry out fraction identification was negligible. In addition, some of the children not eligible for participation in the current investigation because they were identified by pretest performance as capable of fraction identification were informally evaluated on the set partitioning skill. Many of these children did not display set division skill. Not finding many children capable of subset division while at the same time incapable of fraction identification hampered the investigation of hypotheses involving subset division as a prerequisite to fraction identification. In addition, this condition coupled with the informal findings that at least some children in the general population who are capable of fraction identification did not display set partitioning skill raised doubts as to the validity of the
Burke et al. (1977) assertion that subset division is prerequisite to fraction identification.

Further doubt is cast on the assumption of prerequisites with respect to subset division by finding that there was virtually no generalization from fraction identification to subset division. According to Gagne (1977), when a superordinate skill is learned in the absence of prerequisite skill instruction the prerequisite skill(s) will be mastered in the course of superordinate skill learning. Since mastery of subset division did not accompany accurate performance on either one-element or denominator-rule problems, subset division cannot be regarded as a prerequisite to fraction identification according to Gagne's criteria for defining prerequisite skills.

The assumption that subset division should be prerequisite to fraction identification seems entirely reasonable in that set partitioning occurs as part of the fraction identification process. Why then did subset division not emerge as a prerequisite to fraction identification? One possible explanation has to do with differences in the stimulus conditions producing set partitioning behavior in fraction-identification problems and in subset-division problems. In fraction identification, subset division occurs in response to the task of designating a particular fraction. By contrast, in subset division tasks set partitioning occurs in response to the request to divide a set
into a specified number of equal parts. It may be that these two types of stimulus conditions differentially affect the probability that set partitioning activity will take place. Therefore, even though set partitioning appears to be a component of the fraction identification problem, when presented as a task by itself as suggested by Burke et al. (1977) the task may take on additional skills not required in the fraction identification task.

The findings of this study which suggest that set partitioning as advanced by Burke et al. (1977) may not be a prerequisite to fraction identification underscore a point made sometime ago by Glaser and Resnick (1972). These authors suggest that it is important to distinguish between subject matter structure as it is organized by an expert in a particular discipline and subject matter structure as it relates to instruction for a novice in the discipline. At this point in time, the Burke et al. (1977) hierarchy is the major attempt to develop a fraction structure. The Burke et al. hierarchy, however, reflects the organization of the fraction skills as viewed by experts in the field. Although Burke et al. and many of their students (e.g., Heibert and Tonnessen, 1977) have conducted pilot studies related to their fraction identification hierarchy, these pilot studies have in the main addressed issues other than the validation of the hierarchical structure in the purposed fraction organization. The studies have been chiefly concerned with the
strong and weak points of specific research techniques used in the investigation of initial fraction concepts and have not been concerned with validating the relationship between purposed skills.

The principal finding of significance in this study centered around the effect of the types of exemplars used in instruction on rule acquisition. When presented with exemplars compatible with both the one-element and the denominator rules, the vast majority of children extracted the one-element rule rather than the more complex denominator rule. This finding supports the hypothesis advanced in Chapter 1 that when children are exposed to exemplars that reflect more than one rule they will tend to acquire the simplest rule with which the exemplars are consistent. This finding also has significance in that it has implications which run counter to accepted conventional wisdom. One conclusion which has been drawn from hierarchy research over the years is that the content of hierarchies should be taught in sequence for some children, especially for less gifted students, while gifted students may be challenged with more difficult tasks with the assumption that they will learn the prerequisite along the way. Present findings challenge this conclusion. As noted, it appears that it is important to choose stimuli not on the basis of how easy they are but rather on the basis of how unambiguous they are.
In addition to implications concerning the ordering of instruction, results suggesting the extraction of simple rules given ambiguous exemplars have other important implications for instruction. Teachers often attempt to make new concepts clear by providing large numbers of simple examples for children. In so doing, they may inadvertently promote the acquisition of rules that are more limited in scope than the ones intended. Because of this fact, teachers and curriculum developers would do well to take a careful look at the kinds of rules that might be consistent with exemplars used in instruction.

The present findings suggest that exemplars may play an even more important role in determining rule acquisition than verbal instructions accompanying the exemplars. In this connection, it is worth pointing out that children instructed with one-element exemplars tended to extract the one-element rule even though the verbal instructions used in training reflected the denominator rule, that is, included subset partitioning.

The fact that exemplar type may override verbal instruction may pose particularly difficult problems for educators. It would be quite easy for a teacher or instructional-materials developer to be lulled into a false sense of security as to what was being taught if they were to attend mainly to the verbal instructions specifying the concepts to be acquired through instruction. As the present
results show, rule verbalization does not guarantee that the rule verbalized will be the rule that is learned.

The findings for children in the denominator-rule instruction condition indicate that when unambiguous exemplars are provided children can extract a complex rule without prior knowledge of a prerequisite rule. As indicated earlier, the Bergan et al. (1981) study revealed that one-element problems are prerequisite to denominator-rule problems. The present results demonstrated that for large numbers of children it is not necessary to provide instruction in the one-element rule in order to insure denominator-rule acquisition. This finding indicates that appropriate selection of exemplars may make it possible for many students to skip steps in learning a hierarchy. Indeed, the previously mentioned tendency of teachers to afford simple exemplars for complex concepts may foster hierarchical sequencing which is unnecessary.

With respect to the skipping of hierarchical steps, it is particularly important to note that the children who received denominator-rule exemplars not only mastered denominator-rule problems but also were able to generalize the skills they acquired to the simpler one-element rule problems. Thus, nothing was lost by skipping one-element instruction.

It should be noted that not only did the present findings support downward generalization but also they are
consistent with the possibility of upward hierarchical generalization of the type hypothesized by Gagne (1977). While it is true that very few individuals receiving one-element instruction were able to perform denominator-rule problems accurately, it is nonetheless the case that the odds of successful performance on denominator-rule tasks were about two and a half times higher for children who had mastered one-element problems than for children who had not mastered such problems.

Although many children acquired the complex denominator rule without benefit of instruction in the one-element rule, there were some students who did not acquire the denominator rule as a result of instruction with denominator-rule exemplars. The informal analysis of individual difference variables suggested several variables which did not appear to be related to this finding and a single variable which might be associated with it. The individual-difference variables considered to be irrelevant to non-generalization in this particular sample of children included intelligence as assessed by the child's classroom teacher, socioeconomic status, ethnicity and special learning difficulties as delineated by the classroom teacher. A variable which may be related to generalization is the age of the learner. It was found that in the denominator-rule training group 56% of the youngest children, that is, 56% of those children who were six and seven years old, did not show generalization as
measured on the denominator-rule posttest. Two-thirds of these non-generalizers, however, were able to pass the one-element subset. Knowledge which could be gained from further investigations into the possible age-related cognitive variables that might be associated with young children's fraction identification performance would be helpful in planning instructional sequences.

The findings of the investigation of individual-difference variables underscores the need to accommodate individual differences in instruction. More specifically, although most children may be expected to profit directly from denominator rule instruction, there may be children for whom either a delay in fraction identification instruction or one-element exemplars would be appropriate.

The results of the error analysis are consistent with the findings in the modified path analysis related to training group effects on accurate posttesting performance. The fact that the odds of making errors congruent with the one-element rule were markedly higher for children training with one-element exemplars than for children trained with denominator-rule exemplars supports the hypothesis that the one-element rule was not only acquired as a result of one-element exemplar instruction but also that the rule was generalized to conditions in which it did not apply. It is of interest to note that the odds of making an error congruent
with the one-element rule tended to be less than one for children trained with denominator-rule exemplars indicating that these children tended not to make inappropriate generalizations of the one-element rule. This was the case despite the fact that two-thirds of the children trained with the denominator rule and failing the denominator-rule posttest problems were able to pass the one-element posttest.

Protocols obtained from the think-aloud students during training and posttesting suggest the nature of the rules being employed by them and possibly by the other students as well. For example, the think-aloud child taught fraction identification with one-element exemplars counted subsets for each problem on the two posttests. However, on denominator-rule problems in which each subset was composed of two elements the child created and counted aloud subsets containing only one element. Examples of the subset marked by this child were shown in Figure 4 found in Chapter 4. Both verbal and written behaviors of this child would suggest that she equated one element with one subset. This is the base-rule which forms the one-element approach to fraction identification.

The verbalizations of the think-aloud student taught with denominator-rule exemplars suggest that he worked with a conceptualization different from that employed by the child taught with one-element exemplars. For example, the denominator-rule child, during the identification of the fraction one-fifth, marked six squares with an X and then
began to erase some of the marks. When asked why he was making a change in his answer, he pointed to a column of two squares and stated, "It says one-fifth so it is only one of these. Two squares!" There is little doubt that this child was working with the denominator-rule. For him, the total group of elements was equally divided into the required five subsets before he responded to the fraction numerator.

There is also evidence in the protocol for the child taught with denominator rule exemplars to suggest an adaptive production system. For example, when working a one-element problem related to the fraction two-fifths, the student read, "Mark an X on two-thirds of the squares." He then stated, "There are five squares so I have to mark one X, two X." The difference between thirds and five elements was apparently unimportant to the child. This could suggest that he had formed an efficient production to be used with one-element problems. Within this production system, the apparent condition was a single row of elements within a set and the apparent action was to respond to the numerator only. Other verbalizations of this child support the above conceptualization. In general he responded to one-element posttest problems by counting a number of elements consistent with the numerator. For example, in responding to one-third of three triangles, he stated, "One X." and proceeded to the next problem. Responses by this child to denominator-rule problems did not suggest the same production system
although an efficiency move was also implied with respect to these problems. With denominator-rule problems the child's verbalizations clearly suggested set-partitioning relevant to the denominator. For example, when responding to two-thirds of six circles, he stated, "One X, two X. One set. One X, Two X. Two sets." It is the case that the word set is being used incorrectly in these instances, but it is also most probable that the units the child referred to as sets were being treated as specific and equal parts of a larger unit. Efficiency in handling denominator-rule problems was noted in that the child did not draw lines between subsets as had been demonstrated during training. He appeared to cognitively make the division and then mark a numerator-appropriate number of units.

In addition to suggesting the possible rules and production-system changes underlying the fraction identification of the think-aloud children, their verbalizations suggest the possibility of learner confusion. For example, the child taught with one-element exemplars seemed to experience some confusion in keeping with a transition state (Cancelli, Bergan and Taber, 1980). At one point following posttesting the one-element child was asked to explain the meaning of equal parts in a set partitioning problem. She responded that the six equal parts were the lines that divided the boxes. Still within the context of set partitioning, the student was asked what the boxes were. Her reply
was, "What you put the X in for five parts of six." The
verbalizations of the student taught with denominator-rule
exemplars do not indicate confusions of this type.

The major unanswered question in the present inves-
tigation has to do with the problem of determining why some
children receiving denominator-rule instruction mastered
denominator-rule problems and other children receiving such
instruction failed to master the denominator rule. It is
reasonable to assume that there are prerequisite competen-
cies associated with the denominator-rule that were not
identified in this study. Future research might profitably
be directed toward the empirical identification of those
prerequisites.
The following pages contain examples of all testing and training materials employed during the course of the fraction-identification investigation. Each of the materials was color coded to increase stimulus novelty for the learners.
EXHIBIT 1: Pretest used to screen students to determine participation eligibility

School: ___________________________ Name: ___________________________
Grade: ___________________________ Age: ___________________________ Ethnicity: ___________________________

1. Make an X on 1/3 of the circles:
   □ □ □

2. Make an X on 2/3 of the triangles:
   △ △ △

3. Make an X on 2/3 of the squares:
   □ □ □

4. Make an X on 1/3 of the circles:
   □ □ □
EXHIBIT 2: Prerequisite skills test

School: __________________________ Name: __________________________
Grade: ________________________ Age: __________________________ Ethnicity: __________________________

1. Divide this set into 6 equal parts. Draw a circle around each of the 6 parts.

```
△ △ △ △ △ △
△ △ △ △ △ △
△ △ △ △ △ △
```

2. Divide this set into 5 equal parts. Draw a circle around each of the 5 parts.

```
□ □ □ □ □
□ □ □ □ □
□ □ □ □ □
```

3. Divide this set into 4 equal parts. Draw a circle around each of the 4 parts.

```
○ ○ ○ ○
○ ○ ○ ○
○ ○ ○ ○
```

4. Divide this set into 3 equal parts. Draw a circle around each of the 3 parts.

```
△ △ △
△ △ △
△ △ △
```
5. Divide this set into 6 equal parts. Draw a circle around each of the 6 parts.

6. Divide this set into 4 equal parts. Draw a circle around each of the 4 parts.

7. Divide this set into 5 equal parts. Draw a circle around each of the 5 parts.

8. Divide this set into 3 equal parts. Draw a circle around each of the 3 parts.
EXHIBIT 3: Training packet for one-element problems

Mark an X on 3/4 of the triangles:

\[ \triangle \triangle \triangle \triangle \]

Mark an X on 5/6 of the squares:

\[ \square \square \square \square \square \square \]
Mark an X on 1/4 of the circles:

○ ○ ○ ○

Mark an X on 1/6 of the triangles:

△ △ △ △ △ △
Make an X on 5/6 of the squares:

\[
\square \square \square \square \square \square
\]

Make an X on 1/6 of the triangles:

\[
\triangle \triangle \triangle \triangle \triangle
\]
Make an X on 1/4 of the triangles:

△ △ △ △

Make an X on 3/4 of the circles:

○ ○ ○ ○ ○
EXHIBIT 4: Training packet for denominator-rule problems

Mark an X on 3/4 of the triangles:

△ △ △ △ △ △ △ △

Mark an X on 5/6 of the squares:

□ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □
Mark an X on \( \frac{1}{4} \) of the circles:

\[
\begin{array}{cccc}
\bigcirc & \bigcirc & \bigcirc & \bigcirc \\
\bigcirc & \bigcirc & \bigcirc & \bigcirc \\
\bigcirc & \bigcirc & \bigcirc & \bigcirc \\
\end{array}
\]

Mark an X on \( \frac{1}{6} \) of the triangles:

\[
\begin{array}{cccccccc}
\triangle & \triangle & \triangle & \triangle & \triangle & \triangle \\
\triangle & \triangle & \triangle & \triangle & \triangle & \triangle \\
\triangle & \triangle & \triangle & \triangle & \triangle & \triangle \\
\end{array}
\]
Make an X on 5/6 of the squares:

Make an X on 1/6 of the triangles:
Make an X on \( \frac{1}{4} \) of the triangles:

\[
\begin{array}{cccc}
\triangle & \triangle & \triangle & \triangle \\
\triangle & \triangle & \triangle & \triangle \\
\triangle & \triangle & \triangle & \triangle \\
\end{array}
\]

Make an X on \( \frac{3}{4} \) of the circles:

\[
\begin{array}{cccc}
\bigcirc & \bigcirc & \bigcirc & \bigcirc \\
\bigcirc & \bigcirc & \bigcirc & \bigcirc \\
\bigcirc & \bigcirc & \bigcirc & \bigcirc \\
\end{array}
\]
EXHIBIT 5: One-element and Denominator-rule posttest

School: __________________________________ Name: ___________________________
Grade: ____________ Age: ______________ Ethnicity: __________________________

1. Mark an X on 1/3 of the squares:
   □ □ □

2. Mark an X on 1/5 of the circles:
   ○ ○ ○ ○ ○

3. Mark an X on 2/3 of the triangles:
   △ △ △

4. Mark an X on 1/5 of the squares:
   □ □ □ □ □

5. Mark an X on 2/3 of the circles:
   ○ ○ ○

6. Mark an X on 1/3 of the triangles:
   △ △ △
7. Mark an X on 1/3 of the triangles:

\[\triangle \triangle \triangle \triangle \triangle \]

8. Mark an X on 2/5 of the circles:

\[\bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \]

9. Mark an X on 2/5 of the squares:

\[\square \square \square \square \]

10. Mark an X on 1/3 of the circles:

\[\bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \]

11. Mark an X on 2/3 of the triangles:

\[\triangle \triangle \triangle \]

12. Mark an X on 2/5 of the circles:

\[\bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \bigcirc \]
13. Mark an X on 1/5 of the squares:

```
□ □ □ □ □
□ □ □ □ □
```

14. Mark an X on 2/5 of the circles:

```
● ● ● ● ●
```

15. Mark an X on 1/5 of the squares:

```
□ □ □ □ □
```

16. Mark an X on 2/3 of the triangles:

```
△ △ △
△ △ △
```
APPENDIX B

THINK-ALOUD PROTOCOLS
(The presentation of each training fraction was based on the following format. "It says here that we are to mark an X on 3/4 of the triangles. Three-fourths means three parts of four parts. First we need four equal parts. One. Two. Three. Four. Now we need to mark an X on the triangles in three of four parts. One part. Two parts. Three parts.")

1. Mark an X on 3/4 of the triangles. (Subject is reading as teacher presents the task.)
   (T: Now here is one for you that is just like mine. Be sure to think aloud.)

2. Mark an X on 3/4 of the triangles. (Subject is reading the directions.)

3. One. Two. Three. Four. (Subject is counting as she draws lines between subsets.)

4. One. Two. Three. (Subject is counting the X markings as they are drawn.)
   (T: That's three-fourths of the triangles. Exactly right, Kelly. This next one says . . .)

5. Mark an X on 5/6 of the squares. (Subject is reading directions as teacher presents the task.)

6. Five-sixths means five parts of six parts. (Subject is
repeating teacher's words.)

7. One. Two. Three. Four. Five. Six. (Subject is counting along with teacher.)

   (T: Now as you work your problem be sure to think aloud.)

9. Okay. Mark an X on five-sixths of the squares. (Subject is reading the directions.)

10. Okay. I am going to make a line between, I'm going to separate each of the squares for six parts. One. Two. Three. Four. Five. Six lines between each part.

11. Now it says five so I will mark an X on the six parts. One. Two. Three. Four. Five. I have an X in five squares.
   (T: That's very good. Your work is right and you are thinking aloud nicely, Kelly. Here is our next one. . . .)

12. One-fourth of the triangles. (Subject is reading in part with the teacher.)

13. One part of four parts. (Subject is talking along with teacher.)

14. One. Two. Three. Four. (Subject is counting as teacher counts.)

15. One. (Subject is counting as teacher counts.)
   (T: Now here is one for you just like mine.)

17. Now I will mark one.

(T: Perfect. This next one says . . .)

18. Three-fourths means three parts of four parts. (Subject is talking as teacher says the same words.)

19. Four equal parts. (Subject is repeating part of teacher's phrase.)

20. One. Two. Three. Four. (Subject is counting with teacher.)

21. Three. (Subject is repeating the last numerator count made by teacher.)

(T: Here is one for you. Three-fourths.)

22. Make an X on three-fourths of the circles. (Subject is reading the directions.)

23. One. Two. Three. Four. (Subject is counting lines as they are drawn.)

24. One part. Two part. Three parts. (Subject is counting Xs as they are being drawn.)

(T: Exactly right. Let's go on to this next one . . .)

25. . . . on five-sixths of the squares. (Subject partially repeats the directions being stated by teacher.)

26. Five-sixths. (Subject partially repeats the direction being stated by teacher.)

27. Five parts of six parts. (Subject partially repeats
29. Xs on the squares in five parts. (Subject partially repeats teacher's words.)
30. One parts two parts, three parts, four parts. (Subject is repeating teacher.)
   (T: Here is one for you. . . .)
31. Okay. One. Two. Three. Four. Five. Six. (Subject is counting as she draws separation lines.)
32. Now I mark five parts. One. Two. Three. Four. Five. (T: Absolutely right, Kelly. This next one says . . .)
33. One. Two. Three. Four. Five. Six. (Subject is counting with teacher.)
34. One. (Subject is counting with teacher.)
   (T: Here is one for you. One-sixth.)
35. Okay. One. Two. Three. Four. Five. Six. (Subject is counting as she draws separation lines.)
36. Does it matter where I put the mark?
   (T: No.)
37. Okay. One.
   (T: Exactly right. We have just two more to do. Mark an X . . . .)
38. One-fourth means one part of four parts. (Subject is stating the training format along with the teacher.)
39. Four equal parts. (Subject is stating training format along with teacher.) Two. Three. Four. Now we need to mark an X on one of the four parts. Okay. (T: Here is one for you. One-fourth.)

(T: Exactly right. Here's our last one. Mark an X . . .)

41. One-sixth means one part of six parts. (Subject is stating training format along with teacher.) Equal parts. One. Two. Three. Four. Five. Six. Now we need . . . one part.
(T: Here's yours. One-sixth.)

42. One. Two. Three. Four. Five. Six. (Subject is counting as she makes separation lines.)

43. One.
(T: Great. Thank you, Kelly. Now here are some new sets. Some of them are like the ones we have just done. Some of them are different. Be sure to think aloud as you work them. Start here.)


Subject continued in exactly this manner through the entire two posttests. She always read the direction, counted as she made the lines, counted as she marked
the Xs, and concluded each problem by repeating "okay."

Following the full teaching package, the subject was presented with a problem sheet containing denominator-rule exemplars. Two exemplars were demonstrated and then the following conversation occurred.

(T: If I were to ask you to circle each of six equal parts in this set, what would you do?)

1. I would circle that (pointing to a square) and that (pointing to another).

(T: What are the six equal parts?)

2. The lines that divide the boxes.

(T: What are the boxes?)

3. What you put the X in for five parts of six parts.
Student Number 2 (Age 10 years, Grade Four, Taught with Denominator-Rule Exemplars)

(The presentation of each training fraction was based on the following format. "It says here that we are to mark an X on 3/4 of the triangles. Three-fourths means three parts of four parts. First we need four equal parts. One. Two. Three. Four. Now we need to mark an X on the triangles in three of the four parts. One part. Two part. Three parts.")

1. Oh. Now I get it. (This statement was made following the training statement "Now we need to mark the triangles in three of the four parts.")

(T: Now here is one for you. Three-fourths.)

2. One three, two threes, three threes, four threes.

(Subject is counting as he marks the set.) One on the first line, one X, two X, three X, one-fourth. One X, two X, three X, two-fourths. One X, two X, three X, three-fourths.

(T: That's right. Three-fourths. This next one says . . .)

(T: Here's one for you. Five-sixths. Be sure to think aloud as you work.)

two, three. Three parts. One, two, three. Four parts. One, two, three. Five parts.

(T: Very nice. You're absolutely right. Here's another one. This one asks us to . . .)

(T: Here's one for you. One-fourth.)


(T: Okay. Exactly right. Now we are going to do this next one . . .)

(T: Here's one for you. Three-fourths.)


(T: Great. You are learning this quickly.)

6. I thought it was the opposite way. I thought three-fourths was three of the one-fourths. That's how I thought you were supposed to do it.

(T: We can talk about that when you finish. I will explain.)

7. Alright.

(T: Mark an X . . .)

(T: Here's one for you. Five-sixths.)

Five.

(T: This one asks us to . . .)

9. This is easy.

10. One. Two. Three. Four. Five. Six. (Subject is counting with teacher as she marks subsets.)

(T: Here is one for you. One-sixth.)


(T: You're right, Geoff. This next one says mark . . .)

12. One. Two. Three. Four. (Subject is counting with teacher as she marks subsets.)

(T: Here's one for you that's just like mine. One-fourth.)


(T: Good. This is our last one. Mark an X . . .)

(T: Here's one for you. One-sixth.)


(T: That's exactly right. You have marked one-sixth.)

15. Do we have more?

(T: Now we have some new sets. Some of them are like the ones we have just done. Some of them are different. Be sure to think aloud as you work them. Start here.)

16. Mark an X on one-third of the squares. (Subject is
reading directions.)

One.

17. Mark an X on one-fifth of the squares. (Subject is reading directions. He then marked the squares without thinking aloud.)

(T: You did not think aloud as you worked this one.)

18. Okay. One X in two of the circles. Leave the rest empty.

(T: Go ahead with the next one. Be sure to think aloud as you work.)

19. Mark an X on two-thirds of the circles. One X, two X. One set. One X, two X. Two sets.

20. Mark an X on one-fifth of the squares. One X, two X, three X, four X, five X, six X. Oops. (Subject began to erase Xs at this point.)

(T: Why are you correcting that one?)

21. It says one-fifth so it is only one of those. Two squares. (Subject pointed to the first column of elements in the set.)

22. Mark an X on one-third of the triangles. One X.

23. Mark an X on one-third of the triangles. Okay. So that's two Xs. One X. Two X.

24. Mark an X on two-fifths of the circles. So that's one X, two X. One set. One X, two X. Two sets.

25. Mark an X on two-thirds of the squares. (This problem really reads two-fifths of the squares.) There are
five squares so I have to make one X, two X.

26. Mark an X on one-third of the triangles. (The set in this problem actually contains circles and the directions refer to circles.) One X, two X. Two-thirds. No. One-third.

27. Mark an X on two-thirds of the triangles. One X. Two X. Two-thirds of the triangles.

28. Mark an X on two-fifths of the circles. One X. Two X. Three X. Four X. Two. Two. What was it again? Fifths.

29. Mark an X on one-fifth of the circles. One-fifth. One Square.

30. Mark an X on two-fifths of the circles. One X. Two X. Two-fifths.

31. Mark an X on one-fifths of the circles. (The problem contains squares rather than the circles and the directions refer to squares.) One X. One box. One-fifth.

32. Mark an X on two-thirds of the boxes. Triangles. One X. Two X. Oops. That's only one-third. (Two more boxes are then marked).
REFERENCES


Novillis, C. F. An analysis of the fraction concept into a hierarchy of selected subconcepts and the testing of the hierarchical dependence. Journal for Research in Mathematics Education, 1976, 7, 131-144.


---


Watson, J. B. Psychology as the behaviorist views it. *Psychological Review*, 1913, 20, 158-177.


