DOCUMENT RESUME

ED 073 412 CG 007 910

AUTHOR Swanson, James Edward
TITLE The Effects of Number of Positive and Negative Instances, Concept Definition, and Emphasis of Relevant Attributes on the Attainment of Three Environmental Concepts by Sixth-Grade Children: Report from the Conditions of Learning and Instruction Component of Program 1. Technical Report No. 244.
INSTITUTION Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Research and Development Centers Branch.
REPORT NO TR-244
BUREAU NO BR-5-0216
PUB DATE Nov 72
CONTRACT OEC-5-1-154
NOTE 164p.

EDRS PRICE MF-$0.65 HC-$6.58
DESCRIPTORS Concept Formation; *Concept Teaching; Elementary Education; Environment; Environmental Criteria; *Environmental Education; Social Factors

ABSTRACT

Three main studies were conducted to determine the effects of certain independent variables (number of positive and negative instances, presence or absence of a concept definition specifying the defining attributes, and presence or absence of emphasis of relevant attributes) on the learning of environmental concepts. Sixth-grade children who participated in the experiments each read a series of three printed lessons and took four tests dealing with the concepts population, habitat, and community. Five dependent measures were used to assess concept attainment: recognition of new concept instances and non-instances, overgeneralization, undergeneralization, recognition of the concept definition, and knowledge of interrelationships among concepts. Results included: (1) presenting examples and non-examples in a rational teaching set promoted correct classification of unencountered instances; (2) removal of negative instances from the rational teaching set resulted in significant overgeneralization; and (3) providing a concept definition with the rational set of 3 or 4 instances and 3 or 4 non-instances was not more effective in promoting concept learning than providing the rational set alone. (Author/SK)

TECHNICAL REPORT

THE WISCONSIN RESEARCH AND DEVELOPMENT CENTER FOR COGNITIVE LEARNING

The University of Wisconsin Madison, Wisconsin

Technical Report No. 244

THE EFFECTS OF NUMBER OF POSITIVE AND NEGATIVE INSTANCES, CONCEPT DEFINITION, AND EMPHASIS OF RELEVANT ATTRIBUTES ON THE ATTAINMENT OF THREE ENVIRONMENTAL CONCEPTS BY SIXTH-GRADE CHILDREN

Report from the Conditions of Learning and Instruction Component of Program I

By James Edward Swanson

Vernon L. Allen, Frank H. Farley, Herbert J. Klausmeier, and Joel R. Levin
Principal Investigators

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

November 1972

This Technical Report is a master's thesis reporting research supported by the Wisconsin Research and Development Center for Cognitive Learning. Since it has been approved by a University Examining Committee, it has not been reviewed by the Center. It is published by the Center as a record of some of the Center's activities and as a service to the student. The bound original is in The University of Wisconsin Memorial Library.

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.

Center No. C-03 / Contract OE 5-10-154

STATEMENT OF FOCUS

Individually Guided Education (IGE) is a new comprehensive system of elementary education. The following components of the IGE system are in varying stages of development and implementation: a new organization for instruction and related administrative arrangements; a model of instructional programing for the individual student; and curriculum components in prereading, reading, mathematics, motivation, and environmental education. The development of other curriculum components, of a system for managing instruction by computer, and of instructional strategies is needed to complete the system. Continuing programmatic research is required to provide a sound knowledge base for the components under development and for improved second generation components. Finally, systematic implementation is essential so that the products will function properly in the IGE schools.

The Center plans and carries out the research, development, and implementation components of its IGE program in this sequence: (1) identify the needs and delimit the component problem area; (2) assess the possible constraints--financial resources and availability of staff; (3) formulate general plans and specific procedures for solving the problems; (4) secure and allocate human and material resources to carry out the plans; (5) provide for effective communication among personnel and efficient management of activities and resources; and (6) evaluate the effectiveness of each activity and its contribution to the total program and correct any difficulties through feedback mechanisms and appropriate management techniques.

A self-renewing system of elementary education is projected in each participating elementary school, i.e., one which is less dependent on external sources for direction and is more responsive to the needs of the children attending each particular school. In the IGE schools, Center-developed and other curriculum products compatible with the Center's instructional programing model will lead to higher morale and job satisfaction among educational personnel. Each developmental product makes its unique contribution to IGE as it is implemented in the schools. The various research components add to the knowledge of Center practitioners, developers, and theorists.

ACKNOWLEDGEMENTS

I wish to express my genuine gratitude to Dr. Herbert J. Klausmeier, who served as major professor during the preparation and completion of this thesis. It was his thoughtful comments, guiding direction, and concern which made the long hours spent on this project worthwhile and productive. To Dr. Klausmeier I own a special thanks for the encouragement and understanding which he so generously extended to me during my graduate studies.

My gratitutde is also directed to Dr. Joel R. Levin and Dr. Gary A. Davis for their helpful advice as members of my reading and examining committee. I am deeply thankful to Dr. Dorothy A. Frayer, as well, for her generous expense of time and talents during initial planning and research.

I am indebted to the staff of the Wisconsin Research and Development Center for their professional help and advice. Particular thanks are due to the secretaries and duplicating personnel, who spent a great deal of time and effort in the preparation of research materials and the thesis which followed. Special appreciation is extended to Ed Haertel for his expert assistance during data analysis, and to Jayne White who was in charge of keypunching.

The author is also grateful to the personnel of Third Street Elementary School in Evansville, Wisconsin, and to the staff at Oregon Middle School in Oregon, Wisconsin. Special appreciation is given to the respective principals, Mrs. Joyce Lebedoff and Mr. Edward Guziewski, for their fine cooperation and support. Working with the sixth-graders who participated was especially enjoyable.

Finally, I am deeply indebted to my fellow graduate students and officemates, whose thoughtful suggestions and expert advice were invaluable: Barbara Nelson, Gordon Nelson, Kathy Feldman, and Rick Marliave. Their friendship and companionship I will long cherish.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES

ABSTRACT

Three main studies were conducted to determine the effects of certain independent variables (number of positive and negative instances, presence or absence of a concept definition specifying the defining attributes, and presence or absence of emphasis of relevant attributes) on the learning of environmental concepts. Sixth-grade children who participated in the experiments each read a series of three printed lessons and took four tests dealing with the concepts population, habitat, and community. Five dependent measures were used to assess concept attainment: recognition of new concept instances and non-instances, overgeneralization, undergeneralization, recognition of the concept definition, and knowledge of interrelationships among concepts. From the results of the studies it was shown that: (1) presenting examples and non-examples in a rational teaching set promoted correct classification of unencountered instances; (2) removal of negative instances from the rational teaching set resulted in significant overgeneralization; (3) providing a concept definition with the rational set of 3 or 4 instances and 3 or 4 non-instances was not more effective in promoting concept learning than providing the rational set alone; and (4) providing a rational set of instances and non-instances together with concept definition and emphasis of relevant attributes was more effective in promoting recognition of new instances than a presentation of teaching instances plus definition without emphasis.

I

INTRODUCTION

Concept learning is a topic in psychological research which is relevant to both theory and practice. Clark (1971) estimated that in the past three decades over 250 experimental studies on concept attainment have been conducted in the laboratory. Parallel to all this experimental research, however, has been the continuous involvement of classroom teachers and curriculum specialists in the teaching of concepts. According to Clark most teachers present subject-matter content in a rather inconsistent manner, for not only do teachers differ from one another in their methods of presenting a given concept, but the same teacher may present two different concepts in two totally different ways. Understanding the principles of learning concepts, then, seems to be as worthy a goal for the classroom teacher as for the researcher. With researchers and teachers sharing such common interests, one would expect a great deal of facilitative interaction between laboratory and classroom, but all too often this has not been the case. Only within the past decade have there been attempts to extend laboratory-derived principles of learning to the classroom setting.

Since 1961 the Wisconsin Research and Development Center for Cognitive Learning has been engaged in research on concept learning in laboratory and school settings. The present study is part of the programmatic research underway at the R & D Center to investigate variables which influence the learning of subject-matter concepts. Because there are so many different variables which influence concept acquisition, it may be helpful to consider the different kinds of variables in terms of three major groupings, irregardless of the medium which is utilized for instruction (Klausmeier, Davis, Ramsay, Fredrick, and Davies, 1965): instructional variables, learner variables, and concept variables. Within such a framework the present study is focused only on instructional variables which affect the learning of concepts from a printed instructional medium.

Many of the earlier studies on concept attainment were very specific and well-controlled laboratory experiments in which S typically performed a sorting task on a matrix of geometric figures which varied along a limited number of stimulus dimensions (color, shape, etc.). Bruner earlier had done considerable research on concept learning using such a matrix of geometric forms. The kinds of concepts taught in the classroom, however, are quite unlike the classic Bruner-type concepts, particularly because subject-matter concepts usually have an infinite rather than a specifiable finite number of characteristic features, or attributes. Nearly every subject taught in school can be considered as a collection of related concepts, for as Gagné pointed out, "concept learning constitutes by far the major portion of the learning associated with what is supposed to go on in schools" (Gagné, 1965). In classroom instruction, basic concepts are usually presented first, and as the student builds upon previously mastered concepts he progresses to higher order concepts. Eventually the student is called upon to use his knowledge in problem-solving situations and to perceive interrelationships among the concepts he has studied. One of the most fruitful lines of research to emerge in recent years is the investigation of concept learning in the classroom itself, using subject-matter concepts. Only through studies which explore actual classroom variables and processes can learning research become relevant to the real world of instruction.

BACKGROUND OF THE PROBLEM

Programmatic research at the Wisconsin R & D Center has led to the isolation of many important instructional variables involved in concept learning (Blount, Klausmeier, Johnson, Fredrick, & Ramsay, 1967; Frayer, 1970; Frayer & Klausmeier, 1971; Scott, 1970; Smuckler, 1967). Some of the specific instructional variables which have been studied include: number of examples and non-examples, ratio of positive to negative examples, variety of positive and negative examples, sequence of examples, emphasis of relevant attribute values, instructions to recall previously learned subconcepts, etc.

A great deal of attention has been given by various researchers to investigating the role of positive and negative instances in the acquisition of concepts. In 1969 Markle and Tiemann first introduced a set of instructional materials and slides explaining how to analyze and teach concepts in such a way as to avoid certain common errors in classification. The first step in teaching a concept, as suggested by Markle and Tiemann, is to analyze the concept in terms of its relevant and irrelevant attributes. Then a rational set of teaching examples and non-examples is prepared by choosing enough examples to cover the range of irrelevant attribute values and enough non-examples to exclude each of the relevant attributes. If, for example, a teacher wished to present the concept insect to her science class, she might first analyze the concept in a manner similar to the following:

Concept: Insect

Criterial (Relevant) Attributes
1. invertebrate
2. six jointed legs
3. one pair of antennae
4. body parts divided into head, thorax, and abdomen

Irrelevant Attributes
1. size of legs
2. type of legs, e.g., jointed
3. internal or external skeleton
4. body shape, e.g., elongated, stubby
5. wing size (or absence of wings)
6. mode of locomotion, e.g., flying, crawling
7. color (overall and of body parts)

Since Markle and Tiemann prescribe enough teaching examples to cover the range of irrelevant attribute values, a good set of teaching examples for insect would include invertebrates with six jointed legs, one pair of antennae, and three body parts, which differed from each other in size of legs, type of legs, body shape, etc. In keeping with the prescription, a good set of negative examples would also be presented, and for the concept insect at least four non-examples are needed because there are four relevant (defining) attributes to be excluded one at a time.

Markle and Tiemann (1969) not only devised a set of prescriptions for selecting examples and non-examples to use in instruction, but went on to propose a behavioral measure of what it means to "really understand" a concept. They affirmed that simply restating a definition of the concept which was presented during instruction does not measure understanding. Rather, they measure concept mastery in terms of a student's ability to classify previously unencountered instances as either examples or non-examples of the concept. The student who really understands a concept will be able to generalize to new instances what he learned during instruction. Markle and Tiemann are as interested in the kinds of errors which are made in classification as they are in correct categorization, and it is this concern which has led to some interesting hypotheses about the role of positive and negative instances in concept acquisition. According to

Markle and Tiemann, the function of positive instances is to broaden the student's conceptual boundaries and assist him in generalizing to new instances of the concept. If not enough positive teaching examples are presented during instruction, the hypothesized consequence is an error called undergeneralization, in which the student does not identify all of the new examples because his conceptual boundary is too narrow. On the other hand, negative instances are equally important in preventing the opposite kind of classification error, overgeneralization. When not enough negative teaching examples are presented during instruction, the student's conceptual boundaries are not sufficiently limited for him to discriminate certain instances as non-members of the concept class. Overgeneralization is a classification error in which too many instances are identified as examples because of the student's failure to discriminate properly.

Several researchers at Brigham Young University (Tennyson, Wooley, and Merrill) have been gathering empirical evidence on the Markle-Tiemann prescriptions, and their findings support the position that both examples and non-examples are important in concept teaching. In a study involving the concept trochaic meter (Tennyson, Woolley, & Merrill, 1972), independent variables designated as "probability, matching, and divergency" were manipulated to produce certain specified classification behaviors, including correct classification, overgeneralization, and undergeneralization. Tennyson *et al.* first administered a test in which naive students were asked to classify a large number of examples and non-examples of *trochaic meter* on the basis of the concept definition alone. Some of the more obvious examples and non-examples were easily identified by a large proportion of the students and those were called high-probability instances, whereas other examples and non-examples were more infrequently identified and were termed low-probability instances. The independent variables of matching and divergency referred to how closely examples and non-examples resembled each other in terms of their component attributes. Markle and Tiemann (1969) proposed that restricting the range of teaching examples would cause a student to undergeneralize, or to fail to include all true examples of the concept when classifying new instances. Tennyson *et al.* tested the assumption of Markle and Tiemann by giving students a restricted range of examples consisting of the concept definition and low-probability (subtle) non-examples, but very high-probability (obvious) non-examples. The poor selection of teaching examples resulted in significant undergeneralization, just as hypothesized. In the same study using *trochaic meter* Tennyson *et al.* presented other students with a concept definition and a full range of high- and low-probability examples (obvious as well as subtle) but only very high-probability (obvious) non-examples. This experimental manipulation provided a very poor selection of teaching non-examples, which according to Markle and Tiemann would cause the student to overgeneralize, or to mistakenly identify new non-examples as members of the concept class. The Tennyson et al. results confirmed that a poor selection of teaching non-examples will result in overgeneralization, just as Markle and Tiemann hypothesized.

One deficiency apparent in some of the past concept learning studies is the limited scope of the research implications. Markle and Tiemann have proposed how to teach concepts, but have failed to indicate for which kinds of concepts their model is designed. Tennyson, Woolley, and Merrill (1972) have used concepts such as trochaic meter which are appropriate for college-age Ss, but have not explained why their research findings necessarily have implications for concept learning at the preschool or primary level. Not all concepts are equally complex, and not all concepts are learned in the same way. In discussing the results of their studies, few researchers have attempted to relate their findings to a broader framework, or one which accounts for the level of abstractness and complexity of the concept itself.

To provide a conceptual framework for the research being done on concept learning, Klausmeier (1971) formulated a descriptive model of the cognitive operations involved in acquiring concepts at various levels of abstractness and complexity. The four parts of the model are four levels in the attainment of any concept at successively higher degrees of abstractness and inclusiveness: concrete, identity, rudimentary classificatory, and formal. The process of acquiring a concept usually involves a progression from concrete to formal levels, but this is not always the case, for some concepts may not first be attained until the formal (or highest) level. Klausmeier's operational descriptions of the four levels of concept attainment are as follows:

> Attainment of a concept at the concrete level is inferred when the individual cognizes an object that he has experienced on a prior occasion.
>
> Attainment of a concept at the identity level is inferred when the individual cognizes an object as the same one previously encountered when observed from a different perspective or sensed in a different modality, such as hearing or seeing.
>
> Attainment of a concept at the classificatory level is inferred when the individual responds to at least two different instances of the same class as equivalent even though he cannot name the attributes common to them.
>
> Attainment of a concept at the formal level is inferred when the individual with normal language development can accurately designate certain objects or events as belonging to the same set and others as not belonging to the set, can give the name of the concept, and can name its intrinsic or societally accepted defining attributes.

In relation to the model of cognitive operations, organismic variables (the learner), stimulus variables (the concept), and environmental variables (the instructions) are all operative at each of the four progressive levels. The nature of the instructional variables under investigation in the present study are essentially those which are subsumed under formal concept learning in the model proposed by Klausmeier. The model itself is helpful in relating seemingly diverse studies on different levels of abstractness and complexity, and it seems to offer a more integrated frame of reference than other more limited and simplistic models.

PURPOSE

The present research is concerned solely with instructional variables presumed to influence concept learning at the formal level. The dependent variables selected to measure concept acquisition parallel the kinds of abilities which are tapped at the formal level. Specifically, the dependent measures in this study are: recognition of new concept instances, knowledge of concept definition, knowledge of interrelationships among concepts, and the classification errors of overgeneralization and undergeneralization.

Three main studies are intended to investigate the effects of certain independent variables (number of positive and negative instances, concept definition, and emphasis of defining or relevant attributes) on the attainment of environmental concepts. The specific questions to be answered are:

1. What are the effects of presenting selected numbers of positive and negative instances in the absence of a concept definition?
2. What are the effects of presenting selected numbers of positive and negative instances in the presence of a concept definition?
3. When presenting a Markle-Tiemann number of positive and negative instances, what are the effects of defining the concept and emphasizing the relevant (criterial) attributes?

HYPOTHESES

1. It is hypothesized that the presentation of a Markle-Tiemann number of examples and non-examples in the absence of a definition will result in better performance on the dependent measures than the presentation of selected numbers of examples alone. It is also hypothesized that the exclusion of non-examples will result in significantly more over-generalization.
2. It is hypothesized that when a definition of the concept is provided, there will be no significant differences in performance among treatments, whether a Markle-Tiemann number of examples and non-examples or selected numbers of examples alone are presented.

3. It is hypothesized that when the full Markle-Tiemann set of examples and non-examples is presented, concept definition and emphasis of relevant attributes will each facilitate concept learning performance on the dependent measures.

SUBJECTS

There were 97 Ss in Main Study I, 102 Ss in Main Study II, and 114 Ss in Main Study III. All three main studies were conducted in predominantly rural Midwestern schools with sixth-grade children.

METHOD

Three lessons from the content-area of environmental education were prepared. Each lesson dealt with one of the following concepts: population, habitat, and community. Four tests were developed: a test on population, a test on habitat, a test on community, and a final test over all three concepts. Test items were designed to assess recognition of new concept instances, knowledge of concept definition, and knowledge of interrelationships among concepts.

Ss read each of the lessons and took the appropriate tests, involving a total of about one hour. Approximately two weeks later the same tests were administered as a retention measure.

Statistical comparisons of mean scores on the various dependent measures were made, and the ordered hypotheses for each main study were tested. Tukey pairwise comparisons of mean differences were performed on variables not included in the specified hypotheses. Univariate analyses of variance were also done on various test scores.

SIGNIFICANCE OF THE STUDY

By studying the effects of certain stimulus and instructional variables on the learning of actual subject-matter concepts, it is possible to make more realistic assumptions about the classroom learning process. In this respect, the present study has implications for the classroom teacher who is searching for the "best" way to present a particular subject-matter concept. Insofar as the variables explored in this and other studies are common to many of the usual school subject matters, accumulated research may eventually lead to a formulation of prescriptive guidelines for the teaching of classroom concepts.

The present study provides empirical support for the kinds of prescriptions which have been made by Markle and Tiemann as well as Tennyson, *et al*. The various patterns of correct and incorrect classification which have been studied here may have great utility in making preparation of materials for future studies more efficient. In terms of development of materials for the classroom, the results of this study may have implications for the preparation of textbooks and other printed instructional materials. Insofar as the instructional variables in this study are representative of variables which affect the learning of formal concepts, the present research also has implications for extension and verification of the model of cognitive operations proposed by Klausmeier. Finally, the powerful nature of variables such as the kind investigated in the present study may lead to further hypotheses, prescriptions, and validations concerning the acquisition of concepts through printed instructional materials.

REVIEW OF RELATED LITERATURE

BASIC TERMINOLOGY

Concept learning literature is replete with terms possessing very specialized meanings, and for that reason a brief explanation of some commonly used terms may prove beneficial at the outset. In The Psychology of Thinking, Bourne *et al*. define a concept as "any describable regularity of real or imagined events or objects," and state that "to learn a concept is to acquire an understanding of a formerly unrecognized regularity" (p. 177). Learning a concept, therefore, is actually learning to categorize in a certain way according to certain describable regularities. Another more commonly used label for the "regularities" or "features" or "characteristics" of objects and events is the term attributes. Bruner prefers to think of attributes as a set of "cues" which signal the identity of an object (p. 25). Bourne *et al*. refer to attributes as "discriminable characteristics", but point out that the attributes of a concept may not always be clearly specifiable in terms of physical stimuli (p. 179). The following three distinctions are usually made concerning attributes:

1. Attributes which enter into the definition of a concept, or those discriminable features of a concept which define

the concept class, are termed <u>defining</u> or <u>criterial</u> attributes. The term <u>relevant attribute</u> is frequently used to mean the same as "defining attribute", as in the present study when emphasis of relevant attributes actually refers to those attributes which are included in the definition of the concept.

For example: The defining (criterial, relevant) attribute of the concept <u>volcano</u> is "formation by molten rock pushing up through a hole in the earth's crust".

2. There are other attributes which are still <u>relevant</u> but not criterial for identification of a concept, and as such are not included in a concept definition.

   For example: Other attributes relevant to <u>volcano</u> but not criterial are "being landforms" and "protruding above the earth's surface".

3. Attributes which vary among the members of a concept class (features or characteristics which vary from example to example without altering class membership) are called <u>irrelevant attributes</u>.

   For example: Such attributes as "height", "color of surface", "geographical location", etc. have values which differ from volcano to volcano without changing the probability of being called "volcano", so these attributes are irrelevant.

The relevant or defining attributes of a concept are crucial to the distinction between positive and negative examples: a concept example (or positive instance) possesses all of the specified defining attributes of a particular concept, while a non-example (or negative instance) lacks one or more of the specified defining attributes of the concept. A further distinction involving irrelevant attributes is made by researchers such as Tennyson et al., who use the term "matched" to refer to examples and non-examples which share irrelevant attributes as similar as possible. A relationship between two examples (or positive instances) is termed "divergent" when the irrelevant attributes of the instances are as different as possible, and "convergent" when the irrelevant attributes are as similar as possible.

Two other terms commonly used in concept learning literature are generalization and discrimination. Generalization is the act of grouping or classifying objects or events according to their observable similarities, as when the student applies a certain concept label to examples he has not previously encountered. Markle and Tiemann (1969) define generalization behaviorally as "making the same response (giving the same name) to a new example which differs, in some way, from previously met examples". Discrimination, on the other hand, is the act of separating or distinguishing instances of a certain concept as non-members of the concept class. Behaviorally, discrimination is defined by Markle and Tiemann as "making a different response (using a different label) for a non-example which shares some properties with previous examples".

According to Mechner (1969), "the basic procedure for forming concepts. . .involves teaching the student to generalize within classes and to discriminate between classes" (p. 462). Success in categorizing, then, is a function of both generalization and discrimination, which has led Markle and Tiemann to develop a procedure for teaching the student to generalize and discriminate. They have proposed that to teach generalization, the student must be presented with enough positive instances (examples) to cover the full range of irrelevant attribute values associated with the concept. In similar fashion, to teach discrimination the student must be presented with enough negative instances (non-examples) to systematically exclude each of the relevant (defining) attributes of the concept. Markle and Tiemann refer to the recommended number of examples and non-examples as a rational set of teaching instances.

INSTRUCTIONAL VARIABLES

The major portion of this chapter will be devoted to a review of the literature on certain instructional variables which influence concept learning from textual materials. The four sections to follow will deal with type of instances presented, number of instances presented, presence or absence of definition, and emphasis of relevant (defining) attributes. There is a dearth of literature on the effects of definition and emphasis, so the latter two instructional variables will receive less attention.

<u>Type of Instances</u>

The role of type of instances (positive or negative) in concept learning has long been disputed. Smoke (1933) found no difference in performance between Ss using positive instances and Ss using both positive and negative instances, concluding that negative instances did not retard learning. The vast majority of laboratory studies on concept identification have shown that performance is best when Ss are presented with all positive instances, as opposed to either a mixture of positive and negative instances or negative instances alone (Hovland & Weiss, 1953; Huttenlocher, 1962, 1964; Olson, 1963; Smuckler, 1967). D. Cecil Clark (1971), in reviewing much of the literature on conjunctive concept attainment, has concluded that the bulk of research (20 studies out of 25) indicates a sequence of all positive instances to be more effective than a mixed sequence of both positive and negative instances or a sequence of negative instances alone. Laboratory studies dealing with the transfer of concept learning (either to new instances or to new concepts) have been contradictory as to whether all positive instances or a mixture of positive and negative instances is more facilitative in transfer learning (Fryatt & Tulving, 1963; Masilela, 1964; Smuckler, 1967). Despite the inconclusiveness of various laboratory studies on the role of non-examples in concept learning, a strong case can be built for negative instances on purely rational grounds. Markle and Tiemann view positive and negative instances as complimentary to each other in function, whereby <u>both</u> positive and negative instances are required to prevent certain types of classification errors. If there are no negative instances (or a poor selection of negative instances) the student's conceptual boundaries will not be sufficiently limited for him to discriminate between certain instances that are members of two or more concept classes. The consequence of presenting too few negative instances is the classification error termed overgeneralization in which the student mistakenly identifies some of the negative instances of the concept as positive instances. The opposite kind of classification error is undergeneralization, in which the student is too conservative and fails to identify all true examples of the concept class as a consequence of encountering too few positive instances during instruction. In summary, then, Markle and Tiemann recommend an adequate number of positive instances (examples) to prevent the error of undergeneralization, and an adequate number of negative instances (non-examples) to prevent the error of overgeneralization. Occasionally a student falsely assumes that an irrelevant attribute is relevant, which results in a combination of undergeneralization and overgeneralization termed "misconception": the student identifies some non-examples as examples, and some examples as non-examples. If a student can correctly generalize and discriminate when confronted with previously unencountered instances, Markle and Tiemann conclude that the student has "learned" the concept.

Merrill (1971) postulated a theory for teaching concepts very similar to that of Markle and Tiemann, which calls for concept examples differing widely in irrelevant attributes (for generalization) and concept non-examples with irrelevant attributes resembling those of the given examples (for discrimination). Stated another way, Merrill recommends teaching examples which are "divergent" (irrelevant attributes as different as possible) and "matched" with respect to non-examples (irrelevant attributes of the examples and non-examples as similar as possible). A study by Tennyson, Woolley, and Merrill (1972) investigated various patterns of classification as a function of the kinds of instances which were presented during instruction. Seventy-six college Ss were randomly assigned four versions of lessons on the concept trochaic meter (consisting of sixteen positive and negative instances), and took an acquisition test consisting of 30 previously unencountered instances of the concept. Three independent variables manipulated in the study were probability, divergency, and matching. Probability referred to the difficulty of the concept instances, and was determined by computing the percentage of students in a separate sample who correctly classified each instance, given only a definition of the concept. High-probability instances were ones which were identified by 60% or more of the sample (they were obvious), whereas low-probability instances were ones which were identified by 30% or less of the sample (they were subtle). Tennyson et al. made the following hypotheses for the four combinations of positive and negative instances:

1. Correct classification will occur: if both high- and low-probability instances are presented, if examples differ widely in irrelevant attributes, and if non-examples and examples share similar irrelevant attributes.
2. Overgeneralization will occur: if only low-probability instances are presented, if examples differ widely in irrelevant attributes, and if non-examples and examples are very different in irrelevant attributes.
3. Undergeneralization will occur: if only high-probability instances are presented, if examples differ widely in irrelevant attributes, and if non-examples and examples share similar irrelevant attributes.
4. Misconception will occur: if both high- and low-probability instances are presented, if examples are similar to each other in irrelevant attributes, and if non-examples are very different in irrelevant attributes.

Each of the hypotheses above were confirmed ($p < .01$) in the Tennyson *et al*. study. The significance of the findings becomes apparent when each of the treatment conditions is analyzed in terms of the recommendations of Markle and Tiemann as well as Merrill.

1. The first condition was optimal in terms of positive and negative examples, and resulted in correct classification.
2. The second condition presented a poor selection of negative examples (because non-examples and examples

had irrelevant attributes as different as possible), and resulted in overgeneralization.

3. The third condition presented a poor selection of positive examples (because all the examples were obvious and did not expose Ss to the extremes of the concept class), and resulted in undergeneralization.
4. The fourth condition presented poor selections of both positive and negative examples, and resulted in misconception.

Tennyson, Wooley, and Merrill's data support the position that both examples and non-examples are important for effective concept teaching: a wide range of examples prevents undergeneralization, while a good selection of non-examples prevents overgeneralization.

Tennyson (1971) provided substantial empirical support for the importance of negative instances in two experiments using the concept adverb. The first experiment was a direct replication of the Tennyson et al. (1972) study, in which seventh-graders were told the definition (relevant attributes) of the concept and presented with one of the four combinations of instances used in Tennyson et al. All of the treatment conditions in Experiment 1 had non-examples, and the Ss responded on the posttest exactly as hypothesized in terms of correct classification, overgeneralization, undergeneralization, and misconception. Experiment 2 was designed to test the independent variable of negative instances by removing non-exemplars from each of the four conditions. Based on research findings dealing with negative instances in concept attainment literature (which indicate that human Ss cannot make use of negative instances), " anyson hypothesized that Ss would respond on the posttest exactly as in Experiment 1. Tennyson discovered, however, that on the adverb task used in his study, Ss in Experiment 2 (without negative instances) responded randomly on the posttest. His interpretation was that Ss completely failed to acquire the concept adverb when presented with positive instances alone. Tennyson concluded that negative instances are important in concept acquisition because they force the S to concentrate on the relevant attribute(s) when presented with a matched relationship of examples and non-examples. (In a matched relationship the examples and non-examples are as similar as possible in their irrelevant attributes, so the only differences are the relevant attributes.)

In summary, there is strong evidence (Tennyson, 1971) that contrary to the literature which discounts the value of negative instances in concept attainment, negative instances do exert a unique function in preventing the classification error of overgeneralization, and in forcing the student to "focus in" on the relevant attributes of a concept.

Number of Instances

Studies which have tried to determine the optimal number of instances for teaching concepts have been largely inconclusive as to whether a small number or a large number of instances is more facilitative (Amster, 1966; Callentine & Warren, 1955; Morrisett & Hovland, 1959). In part, this may be due to the fact that concepts differ greatly from one another in number of defining attributes and number of irrelevant attributes. Since concept complexity is a function of the number of attributes a concept possesses, it must be apparent that unless studies are equivalent in terms of complexity the results will be inconsistent. For example, if Study A uses a concept with 4 relevant and 4 irrelevant attributes, and Study B uses a different concept with only 1 relevant and 1 irrelevant attribute, the presentation of 8 instances (4 positive, 4 negative) will not constitute similar treatments for the two studies. In effect, Study B would be employing the equivalent of four times as many instances as Study A.

The model for teaching concepts postulated by Markle and Tiemann (1969) prescribes a number of positive and negative instances which is dependent upon the number of attributes which a concept possesses. Again, Markle and Tiemann recommend enough positive instances to cover the full range of irrelevant attributes (divergent) and enough negative instances similar in irrelevant attributes to the exemplars (matched) to systematically exclude each of the relevant attributes. This number of instances constitutes what is termed the rational set of teaching instances, and is relative rather than absolute because it depends upon the number of relevant and irrelevant attributes of the concept to be taught.

Frayer (1970) presented eight versions of programmed lessons dealing with geometric concepts to fourth- and sixth-grade children to determine the effect of number of instances and emphasis of relevant attribute values on concept mastery. Either 4 concept instances (2 positive and 2 negative) or 8 oncept instances (4 positive and 4 negative) were included in the lessons. Since no instances were repeated in her study, an increase in the number of instances from 4 to 8 implied an increase in the variety of instances as well. The Ss studied lessons for four days, and took a multiple-choice test and a completion test which consisted of eleven types of questions related to concept learning (specifically, recognition of attribute examples, attribute names, concept examples and non-examples, concept names, relevant and irrelevant attributes, concept definition, and concept relationships). The results of Frayer's study indicated that increasing the number of instances from 4 (2 positive, 2 negative) to 8 (4 positive, 4 negative) did not significantly affect overall concept mastery for either fourth- or sixth-grade children, but it did significantly improve recognition of concept non-examples for the fourth graders. In Frayer's study, then, the effect of a wider variety of instances (due to an increased number of examples and non-examples) was to facilitate discrimination on tasks involving geometric concepts.

Remstad (1969) presented a series of plane geometry concepts to fifth-grade children with slides and tape-recorded verbal cues, in which one of the independent variables was number of positive instances (6 or 9). Another independent variable consisted of five ratios of positive to negative instances (all positives; 2 positive to 1 negative; 1 positive to 1 negative; 1 positive to 2 negative; or all negatives). The results of the study showed that there was no significant difference in response when the number of positive instances was increased from 6 to 9, but a change from all positive instances to a ratio of two positive to one negative (2:1) produced substantial increments above baseline responses. In discussing the results of his study, Remstad suggested that classroom instruction might benefit from the inclusion of more negative instances and a wider variety of instances among other things.

In summary, studies comparing the effectiveness of various numbers of instances may have failed to produce clear-cut results because the "optimal" number of instances depends in part upon the complexity of the concept itself, or how many relevant and irrelevant attributes it possesses. When increasing the number of positive and negative instances results in an increase in the variety of the instances as well, recognition of negative instances (discrimination) may improve significantly. It does not seem true, however, that simply increasing the number of positive instances will result in better concept learning.

Concept Definition

The effect of providing a concept definition on the attainment of concepts has never been widely studied. A traditional classroom approach to instruction, though not always effective, has been to give a definition of the concept to be taught and to test the students' recall of the definition. Even though ability to restate a definition is not a true measure of concept understanding, there is some evidence that definitions do facilitate concept acquisition. Merrill and Tennyson (1971) conducted an experiment in which the concept trochaic meter was taught to 180 educational psychology students at Brigham Young University. The four independent variables manipulated in the study were:

1. Definition presentation (D), in which the relevant (defining) attributes of trochaic meter were stated.
2. Attribute definition (A), in which each relevant and irrelevant attribute of the concept are defined and the intended meanings of the subconcepts are explained.
3. Exemplar/non-exemplar presentation (E), in which examples and non-examples are presented according to the procedure used by Tennyson et al. (1972).

4. Attribute prompting presentation (P), in which each example is accompanied by a statement of the relevant attributes and why they are relevant, and each non-example is accompanied by a statement and explanation of the relevant attributes which are lacking.

Eight combinations of the four independent variables constituted the treatment conditions (D, DA, E, DE, DAE, EP, DEP, DAEP). Lessons on trochaic meter identical to the Tennyson et al. study in terms of probability, divergency, and matching were read by the Ss, and a post-test was given which required the Ss to generalize to previously unencountered examples and to discriminate previously unencountered non-examples. Merrill and Tennyson hypothesized that treatment groups D (definition only), E (examples/non-examples only), and DE (definition plus examples/non-examples) would all result in overgeneralization because there was insufficient information to promote discrimination. Results of the study confirmed their hypothesis ($p < .01$) and revealed that the least effective conditions were D (definition of relevant attributes alone) and E (presentation of examples and non-examples alone). An important finding of Merrill and Tennyson was that Ss given only the concept definition performed as well as the group receiving a full set of examples and non-examples without definition. This would suggest that defining a concept by specifying its relevant attributes provides the S with a substantial amount of information.

In her study on the learning of geometric concepts by fourth and sixth graders, Frayer (1970) looked at the effects of number of concept instances and emphasis of relevant attribute values in the presence of a definition which specified the relevant attributes of the concept. She found no difference in performance due to the number of concept instances presented (2 positive and 2 negative vs. 4 positive and 4 negative) when a definition of the concept was provided. Although Frayer did not study the effect of number of instances when a definition was not provided, it does not seem unreasonable to suppose that the effect of number of instances is different when a definition is presented than when a definition is absent.

Anderson and Kulhavy (1972) did a study in which college Ss were exposed to one-sentence definitions of unfamiliar concepts and then answered multiple-choice questions requiring them to select concept instances and non-instances. They found that merely giving a definition significantly affected instance recognition, and concluded that people can easily learn concepts from definitions. It is worth noting that Anderson and Kulhavy did not take restatement of the definition as a suitable measure of concept comprehension. They tested for learning on the basis of the Ss' capacity to generalize to previously unencountered examples and to discriminate previously unencountered non-examples of the concept. This criterion is the same as that used by Markle and Tiemann and by Merrill, and greatly diminishes the effect of Ss who respond on the basis of surface information without comprehension of the concept itself.

In summary, although there is not a great volume of literature on the effects of providing concept definitions, there is evidence that definitions which specify the defining attributes of a concept do have a beneficial effect upon concept learning performance, and may somewhat diminish the effectiveness of other instructional variables because of the amount of relevant information which the definition supplies.

Emphasis of Relevant Attributes

In printed instructional materials, emphasis of relevant attributes is achieved by any cue which effectively draws the S's attention to the relevant attributes of the concept. This may involve arrangement of copy on the page, attention-directing symbols such as arrows and boxes, or the use of informative feedback in various types of lesson formats. The 1971 study by Merrill and Tennyson (discussed in the previous section on Concept Definition) included emphasis of relevant attributes as one of the independent variables, termed "attribute prompting". Attribute prompting was achieved by identifying the relevant (defining) attribute in each example and stressing the absence of the relevant attribute in each non-example. It was found that of all the treatment conditions involving various combinations of concept definition (D), attribute definition (A), exemplars/non-exemplars (E), and attribute prompting (P), the most powerful condition was the one which included instances plus concept and attribute definitions with attribute prompting (DAEP).

Error rates on a transfer task were significantly lower for the conditions with attribute prompting than for conditions without the prompting variable.

Frayer (1970) used attention-directing questions to emphasize relevant attributes in a modified linear programming format. Her study on the effects of number of instances and emphasis of relevant attributes on mastery of geometric concepts revealed an increase in overall concept mastery for fourth- and sixth-grade children in the emphasis condition. In addition, recognition and production of attribute names for sixth-graders increased significantly when the relevant attributes were emphasized.

In summary, studies on the effects of emphasizing relevant attributes show a facilitation of concept learning performance when attribute prompting is included in printed instructional materials.

III

DEVELOPMENT OF INSTRUCTIONAL MATERIALS AND TESTS

To research the effects of instructional variables on learning from text, it was necessary to develop a series of three self-instructional printed lessons, each of which dealt with a specific concept in the field of environmental education. Among those considerations which guided selection of the specific concepts were:

1. The concept should be appropriate in difficulty for mastery by an average sixth-grade child.
2. The concept should be one which has at least two relevant attributes and at least two irrelevant attributes.
3. The attributes of the concept must be readily identifiable, and recognition of the attributes should not require a high degree of subject matter knowledge.

The three environmental concepts finally selected were population, habitat, and community, each of which seemed to fulfill the requirements which had been established. Definitions which specified the relevant attributes of the concepts were written, : fined, and subsequently were approved by three specialists in environmental science. A summary of concept definitions and attributes is given in Table 1.

Before generating examples and non-examples for a given concept, the concept was analyzed into its component attributes: relevant (or criterial) attributes are the features or characteristics of an instance which are common to all examples of a concept, and irrelevant attributes are the features or characteristics of an instance which are not shared in common by all members of the concept class. Markle and Tiemann (1969) alternately described an irrelevant attribute as "a property of any particular example which can be varied without changing the example to a non-example." Briefly, positive examples of the three environmental concepts were generated for the present study by providing values for all of the relevant attributes, and varying the values of irrelevant attributes. Negative examples were generated by using irrelevant attributes similar to those of positive examples and excluding each of the relevant attributes, one at a time. This method of constructing examples and non-examples is illustrated for the concept population in Table 2.

Table 2
Systematic Generation of Examples
and Non-Examples for the Concept Population

Examples (Positive Instances)	Rationale*
1. All of the black bears in Yellowstone Park	(1,2,3,4b,5b)
2. All of the banana trees in the Amazon Jungle	(1,2,3,4b,5b)
3. All of the trout in Lake Michigan	(1,2,3,4b,5a)
Non-Examples (Negative Instances)	
1. All of the stars in the Big Dipper	(2,3) Lacks 1
2. All of the bees and butterflies in a meadow	(1,3) Lacks 2
3. Half of the squirrels in a forest	(1,2) Lacks 3

*KEY TO RELEVANT ATTRIBUTES: 1=living things 2=same kind 3=entire group found in a particular place

*KEY TO IRRELEVANT ATTRIBUTES:
4=kingdom of living things
a) plant b) animal
5=geographical location
a)aquatic b) terrestrial

Table 1

Summary of Concept Definitions

and Attributes

Concept	Definition	Relevant Attributes	Irrelevant Attributes
Population	A population is the entire group of living things of the same kind to be found in a pariticular place.	1. must be living things 2. must be of the same kind 3. must be the entire group found in a particular place	1. may belong to either of the two kingdoms of living things: a) plant b) animal 2. may be found in different geographical locations a) aquatic b) terrestrial
Habitat	A habitat is all of the space, or total area, a living thing needs in order to survive.	1. must be inhabited by living things 2. must be the total space, or enough area to provide life support	1. may belong to either of the two kingdoms of living things a) plant b) animal 2. may be found in different geographical locations a) aquatic b) terrestrial
Community	A biological community is several kinds of plants and animals which live together in a particular place and which need each other to stay alive.	1. must contain plants 2. must contain animals 3. must involve interaction among living things (depend upon each other for food, shelter, etc.)	1. may vary in size a) large b) small 2. may be found in different geographical locations a) aquatic b) terrestrial

INSTANCE PROBABILITY ANALYSIS

Purpose

Once a pool of examples and non-examples had been generated for each of the environmental concepts, there arose the problem of which instances to include in the lesson materials. One method of selecting the "best" examples and non-examples would have been the subjective approach in which E simply excludes all instances he feels to be poor. According to Tennyson and Merrill (1971), subjective rating of items has been the usual procedure in all forms of instructional development, but some alternative method of item selection would be much preferred, especially some procedure for empirically rating examples and non-examples. The purpose of the Instance Probability Analysis, therefore, was to provide an empirically based procedure for selecting the "best" (most appropriate) instances prior to the construction of lesson materials. In particular, an Instance Probability Analysis makes it possible for the researcher to choose examples and non-examples which have roughly the same probability of being correctly identified. Further, by determining such an index of obviousness prior to the writing of lessons, it is possible to achieve more control over the extraneous variable of item difficulty.

Subjects

The Instance Probability Analysis was conducted at Third Street Elementary School in the small rural community of Evansville, Wisconsin. A total of 88 Ss from four sixth-grade classes (which comprised the entire sixth-grade population of the school) participated. Home-room teachers indicated that none of the children had mastered the environmental concepts population, habitat, or community which were to be presented.

Tests

Two parallel forms (Form A and Form B) of an Instance Probability Test were developed, each form consisting of 20 population items (10 positive instances and 10 negative instances), 20 habitat items (10 positive instances and 10 negative instances), and 10 community items (5 positive instances and 5 negative instances). A concept definition which specified the relevant attributes of the concept was printed within a box at the top of each page in the test booklets. Ss were instructed to read each definition carefully, and to identify each of the test items as an example or non-example of the concept. To the left of each test item were the words "Yes" and "No"; if S decided that an item was an example of the concept, he circled "Yes", and if he decided that an item was not an example of the concept, he circled "No". A random order of presentation for examples and non-examples was determined by referring to a table of random numbers.

Procedure

Test booklets were stacked for distribution in alternate fashion (Form A, Form B, Form A,. . .). The proctor introduced himself, gave instructions concerning test procedure, and pronounced a list of difficult words printed in the test booklets. Ss were told to read the directions printed on the Sample Page and to mark the answers to three sample items, after which the correct responses were discussed by the proctor and Ss were directed to begin working. (See instructions to students given in Appendix A.) Two proctors were used: E proctored two classes, and a research colleague proctored the other two classes.

Results

The percentage of Ss correctly identifying each item was computed in the following manner: for each of the 50 items in Form A and for each of the 50 items in Form B the percentage of "Yes" responses to a concept example or the percentage of "No" responses to a concept non-example were calculated. A list of all the test items and their respective probabilities for the Instance Probability Analysis comprises Appendix B. It was found that the concept habitat was most frequently correctly identified, with 70.3% of all items marked correctly by the 88 students. Other percentages were 65.25% of all population items identified correctly, and only 61.5% of all community items correctly identified. On the whole, negative instances were not correctly identified as often as positive instances, and this trend held true for each concept computed separately as well as for all three concepts considered together. When percentages for all positive items were computed (across concept) it was discovered that positive items were marked correctly 82% of the time, but negative items were marked correctly only 49.5% of the time. As shown in Table 3, the range of probabilities for negative instances was much more extreme than the range of probabilities for negative instances. It was not uncommon for negative instance probabilities to be far below the .50 level of chance responding.

Table 3

Range of Probabilities for Positive and Negative Instances Used in the Instance Probability Analysis

	Positive Instances Range			Negative Instances Range		
Form A						
Population	.88	to	.65	.84	to	.37
Habitat	.98	to	.65	.95	to	.12
Community	1.00	to	.71	.65	to	.17
Form B						
Population	.84	to	.57	.68	to	.28
Habitat	.97	to	.61	.95	to	.11
Community	.82	to	.66	.51	to	.22

PILOT STUDY

Purpose

Objectives of the Pilot Study were:

1. To determine the level of difficulty of lessons and tests for sixth-grade children.
2. To determine the suitability of the modified linear program format which was used for all three lessons.
3. To collect data on item probabilities to be used in test revision.
4. To determine exact time requirements for lessons and tests.

Subjects

Twenty-eight students from Oregon Middle School in Oregon, Wisconsin served as subjects in the Pilot Study. The school is located in a small, middle-class, rural community in southern Wisconsin. In design, the Oregon Middle School is an open classroom, in which approximately 550 children are separated into four "units" of nearly equal size. Only sixth-grade children were used in this study; however, Ss were selected from Unit A in which children range from ages ten to twelve. The 28 Ss were selected on the basis of convenience of scheduling, and they comprised the population of one sixth-grade social studies class. According to the science instructor, none of the students had mastered the concepts population, habitat, or community.

Materials

Lessons

Based on items from the Instance Probability Analysis, three lessons were developed.

LESSON I. In this lesson the concept population was presented. A definition of the concept was provided, as well as a Markle-Tiemann proportion of instances: 4 examples and 3 non-examples, sequenced +,+,+,+,-,-,-.

LESSON II. The concept habitat was presented. There was a concept definition and a Markle-Tiemann proportion of instances: 4 examples and 2 non-examples, sequenced +,+,+,+,-,-.

LESSON III. The biological concept of community was presented. There was a concept definition and a Markle-Tiemann proportion of instances: 4 examples and 3 non-examples, sequenced +,+,+,+, -,-,-.

All three lessons were written in a modified linear programming format, which required the S to write answers to questions about the relevant attributes, and provided immediate printed feedback along the right-hand margin of the page. Feedback was also provided for three multiple-choice questions and two review questions at the end of each lesson, following the presentation of concept examples and non-examples. Each of three multiple-choice questions required the S to complete a phrase by selecting the correct relevant attribute, as below:

Populations are always made up of ___?___.

a) living things

b) things which are not living

The review questions were written to encourage students to look over the instances again. Since Lesson I dealt with the concept population, the review questions in that lesson were: "How many examples of a population did you find in this lesson?" and "How many groups were not examples of a population?" The readability of all lessons was checked by a reading specialist, and found to be appropriate for sixth graders. The Pilot lessons were very similar to the basic lessons found in Appendix D.

Selection of examples and non-examples for the Pilot Study lessons was based on item statistics from the Instance Probability Analysis. In order to choose examples and non-examples which were roughly equivalent in terms of probabilities, it was necessary to employ the following specific selection rules:

1. The total number of positive instances shall be selected from nearest the median probability, when all positive instances are rank ordered.
2. Each negative instance shall be selected from nearest the median probability, when negative instances are ranked separately by the relevant attribute which is excluded.

A detailed summary of the selection procedure for positive and negative instances of the concepts _population_, _habitat_, and _community_ for the Pilot Study comprises Appendix C.

Tests

Following selection of examples and non-examples for all three concept lessons, the remaining instances in the pool were used in construction of three tests to measure concept acquisition (included in Appendix E).

TEST 1. _Part I_, Identification: _S_s were required to identify twenty-five instances as examples or non-examples of the concept _population_.
Part II, Multiple-choice: _S_s were required to select the correct definition of _population_.

TEST 2. _Part I_, Identification: _S_s were required to identify twenty-five instances as examples or non-examples of the concept _habitat_.
Part II, Multiple-choice: _S_s were required to select the correct definition of _habitat_.

TEST 3. _Part I_, Identification: _S_s were required to identify ten instances as examples or non-examples of the concept _community_.
Part II, Multiple-choice: _S_s were required to select the correct definition of _community_.

In addition to the three tests already described, a fourth test was constructed to measure knowledge of interrelationships among concepts and knowledge of concept definitions (defining attributes).

TEST 4. Part I, True-False: Ss were required to answer five true-false items which dealt with interrelationships.

Part II, Multiple-choice: Ss were required to complete four multiple-choice items which dealt with interrelationships.

Part III, Matching: Ss were required to match concept labels with the appropriate concept definitions for population, habitat, and community.

Procedure

All Ss received the same lesson materials. E explained that the purpose of the lessons and tests was not student evaluation, but evaluation of the instructional materials. The use of cardboard strips to conceal answers printed in the lesson booklets (along the right-hand margins) was demonstrated, E pronounced a list of difficult words, and the students were instructed to begin. When S had completed a lesson, he raised his hand, and the proctor collected his lesson and handed him a test. Ss were encouraged to write comments on the back of each test concerning lesson and test difficulty, interest level, ambiguities, etc. After all Ss had completed Test 1 the tests were collected and Lesson 2 was distributed. The complete sequence of lessons and tests was: Lesson I (Populations), Test 1 (Populations), Lesson II (Habitats), Test 2 (Habitats), Lesson III (Communities), Test 3 (Communities), Test 4 (Final Test).

Results

Probabilities were calculated for every test item used in the Pilot Study. A summary table of probabilities for Part I items requiring the S to identify examples and non-examples comprises Table 4. The ranges of possibilities for positive and negative examples were less extreme than in the Instance Probability Analysis. Mean probability scores were generally high but not at ceiling level.

Table 4

Probabilities for Part I Test Items (Pilot Study)

	Positive Instances		Negative Instances	
Test	Range	Mean	Range	Mean
Populations (Test 1)	1.00 - .93	.96	1.00 - .43	.76
Habitats (Test 2)	1.00 - .86	.95	1.00 - .75	.89
Communities (Test 3)	1.00 - .89	.94	.96 - .75	.88

In Part II of Tests 1, 2, and 3 which required the S to select correct concept definitions, 93% of the Ss selected the correct definitions for population and community, while only 68% selected the correct definition for the concept habitat. The depressed probability for this last concept was apparently due to an ambiguous phrase, which was modified for the main studies.

Results of Test 4 (Final Test) were as follows:

Part I: A mean probability of .86 was obtained for true-false items testing concept interrelationships. An especially ambiguous item (p=.54) was excluded from the main studies.

Part II: A mean probability of .72 was obtained for multiple-choice items testing concept interrelationships.

Part III: On the matching section, probabilities for the correct definitions were .79 (Community), .57 (Population), and .39 (Habitat). Because of the unusually low probability for habitat, an especially ambiguous definition of that concept was modified for the main studies, and directions were clarified by adding "There is only one correct definition for each word!"

Written comments provided by the students generally indicated approval of the modified linear programming format, enjoyment of the lessons, and satisfactory level of difficulty. Though roughly 25% of the Ss suggested that the materials were too simple for sixth graders, all lesson and test materia (with slight modifications) were judged to be suitable for use at the sixth-grade level in the three main studies.

IV

METHOD

To investigate the effects of certain instructional variables on the immediate acquisition and retention of concepts in environmental education at the elementary grade level, three main studies were conducted. Each of the main studies was directed toward one of these specific research questions:

1. What are the effects of presenting selected numbers of positive and negative instances in the absence of a concept definition?
2. What are the effects of presenting selected numbers of positive and negative instances in the presence of a concept definition?
3. When presenting a Markle-Tiemann number of positive and negative instances, what are the effects of defining the concept and emphasizing the relevant (criterial) attributes?

MAIN STUDY I

The first main study investigated the effects of number and type of concept instances on the acquisition of environmental concepts in the absence of a definition.

Subjects

The Ss for Main Study I were students at the same elementary school in which the Pilot Study was conducted. The initial sample consisted of 111 Ss from two "units" of sixth graders at Oregon School in Oregon, Wisconsin, but due to absences at the time of retention testing the sample size used in data analysis was 97. Ss in this study plus those who were involved in the Pilot Study constituted the entire sixth-grade population of one unit (Unit B) and half of a second unit (Unit A) at Oregon Middle School. Children from Unit A were selected on the basis of convenience of class scheduling, and included those students who were in the same social studies section. All those students in this unit who had been exposed to the Pilot materials were later excluded from the main study. Ss in the Pilot group were not told that identical materials would later be presented to fellow sixth graders, and there was a three-week interval between the Pilot and Main Study I, so that contamination of the Main Study sample was judged to be negligible. According to the teachers, students had not been formally exposed to any of the three environmental concepts population, habitat, or community.

Materials

Lessons

For the purposes of this experiment all lesson materials used in the Pilot Study (with modifications described in Chapter 3) were changed in the following ways:

1. No definition of the concept was provided.
2. No emphasis of relevant attributes was provided (the modified linear programming format was not used).

Three versions of each of the concept lessons were prepared, as shown in Table 5. In addition, three placebo lessons were developed: Lesson I, Conservation; Lesson II, Transportation; and Lesson III, Invention. The placebo lessons were condensed versions of articles in The World Book Encyclopedia (1969 edition).

Tests

Test 1 (Populations), Test 2 (Habitats), Test 3 (Communities), and Test 4 (Final Test) were the same tests which were administered in the Pilot Study (with modifications as noted in Chapter 3). The long-term Retention Test was composed of Tests 1-4 combined into a single booklet.

Procedure

Ss were randomly assigned within reading level to one of four experimental conditions. All lessons were pre-packaged, with the S's name and treatment group printed on the envelope. Table 6 shows the lesson and test sequence for each of the treatment groups.

Table 5

Content of Lessons

for Main Study I

Concept	Version	Content
Population	Form S-101	4 + instances* and 3 - instances*
	Form S-102	4 + instances*
	Form S-103	2 + instances
Habitat	Form S-201	4 + instances* and 3 - instances*
	Form S-202	4 + instances*
	Form S-203	2 + instances
Community	Form S-301	4 + instances* and 3 - instances*
	Form S-302	4 + instances*
	Form S-303	2 + instances

*Indicates the number of instances which is recommended by Markle and Tiemann.

Table 6

Lesson and Test Sequence

for Each Group in Main Studies

Treatment group	Day 1							Day 17[a] Day 12[b] Day 15[c]
1	LESSON I (Population) Form S-101	TEST 1	LESSON II (Habitat) Form S-201	TEST 2	LESSON III (Community) Form S-301	TEST 3	TEST 4	RETENTION TEST (Tests 1-4)
2	LESSON I (Population) Form S-102	TEST 1	LESSON II (Habitat) Form S-202	TEST 2	LESSON III (Community) Form S-302	TEST 3	TEST 4	RETENTION TEST (Tests 1-4)
3	LESSON I (Population) Form S-103	TEST 1	LESSON II (Habitat) Form S-203	TEST 2	LESSON III (Community) Form S-303	TEST 3	TEST 4	RETENTION TEST (Tests 1-4)
4	PLACEBO (Conservation)	TEST 1	PLACEBO (Transportation)	TEST 2	PLACEBO (Invention)	TEST 3	TEST 4	RETENTION TEST (Tests 1-4)

[a] Main Study I
[b] Main Study II
[c] Main Study III

The proctor distributed lesson packets and gave instructions concerning procedure to be followed in completing lessons and tests. A list of difficult words was pronounced, questions of a general nature were answered, and students were told to begin work. When S had finished reading a lesson he raised his hand, and the proctor collected his lesson and handed him a test. After S had completed tests 1-3 the proctor collected all three tests and S worked quietly on an assignment until all students were finished, at which time the final tests were distributed.

Students were seated in three clusters to keep group size manageable. Five proctors were used altogether: E and two research colleagues familiar with the study were in charge of the clusters, and two Center employees assisted with the proctoring. Directions to the students were essentially the same as directions for the Pilot Study (Appendix A).

Design

The experimental design for Main Study I was a 3 x 4 randomized block design with three levels of reading achievement (high, medium, and low) and four treatment groups, as diagramed in Table 7. This design was employed for all three main studies.

Previous reading achievement was based on STEP scores (Sequential Tests of Educational Progress, Series II). By blocking on previous level of reading achievement, an equal number of Ss was randomly assigned to each of the four treatment conditions. A summary of the treatment conditions for Main Study I is provided in Table 8, and the number of subjects in each cell is given in Table 9. Independent variables in this experiment were number of concept instances presented and type of concept instances presented (in the absence of a concept definition).

Table 7

Experimental Design

for Main Studies I, II, and III

Previous Reading Achievement Level	Treatment Group			
	Condition 1	Condition 2	Condition 3	Condition 4 (Placebo)
High				
Medium				
Low				

(Note - Arrows refer to randon assignment)

Table 8

Summary of Treatment Conditions

for Main Study I

Treatment	Independent Variable			
	Number of Instances		Definition	Emphasis
	Positive	Negative		
Cond. 1	Markle-Tiemann*	Markle-Tiemann	no	no
Cond. 2	Markle-Tiemann	none	no	no
Cond. 3	2	none	no	no
Cond. 4	Placebo lessons			

* According to the Markle-Tiemann paradigm, the specific number of positive and negative instances depends upon the number of relevant and irrelevant attributes which a concept possesses.

Table 9

Number of Subjects By Cell

for Main Study I

Previous Reading Achievement Level*	Treatment Condition				Totals Across Condition
	MT**positives MT negatives	MT positives	2 positives	Control placebo lessons	
High	8	7	8	7	30
Medium	8	9	9	9	35
Low	8	8	9	7	32
Totals Across Reading Level	24	24	26	23	97

* Based on STEP (Sequential Tests of Educational Progress, Series II) scores

** Refers to the rational set of instances as recommended by Markle and Tiemann

Hypotheses

The following specific hypotheses were made for Main Study I (refer to Table 8 for a summary of treatment conditions):

Ia. On the dependent variable "recognition of new concept instances", Condition 1 will result in significantly better performance than Condition 2, which will be significantly better than Condition 3, which will be significantly better than Condition 4 (Control).

b. On the dependent variable "overgeneralization", Condition 1 will result in significantly less overgeneralization than Condition 2.

II. On the dependent variable "knowledge of concept definition", Condition 1 will result in significantly better performance than Condition 2, which will be significantly better than Condition 3, which will be significantly better than Condition 4 (Control).

III. On the dependent variable "knowledge of interrelationships among concepts", Condition 1 will result in significantly better performance than Condition 2, which will be significantly better than Condition 3, which will be significantly better than Condition 4 (Control).

MAIN STUDY II

The second main study was identical to Main Study I in design and methodology except for the inclusion of concept definitions. The purpose of this study was to investigate the effects of number and type of concept instances on the acquisition of environmental concepts in the presence of a definition.

Subjects

Main Study II was conducted in four different schools within a medium-size Midwestern joint school district. The schools were generally small and located in predominantly rural communities. At the beginning of the study there were 103 sixth-grade Ss, but due to one absence at the time of retention testing, the final number of Ss was 102. Teachers at all four schools indicated that their students had not been formally exposed to the concepts population, habitat, or community prior to the experiment.

Materials

Lessons

For the purposes of this experiment all lessons were identical to the lessons in Main Study I, except that concept definitions for population, habitat, and community were inserted. No emphasis of relevant attributes was provided (the modified linear programming format was not used). Lesson content for the treatment groups in Main Study II is described in Table 10. Placebo lessons were identical to those on Conservation, Transportation, and Invention which were used in the previous study.

Tests

Test 1 (Populations), Test 2 (Habitats), Test 3 (Communities), and Test 4 (Final Test) were the same tests which were administered in Main Study I. The long-term Retention Test was composed of Tests 1-4 combined into a single booklet.

Procedure

The procedure for this study was identical to the procedure outlined for Main Study I. Consult Table 6 for the schedule of lessons and tests.

E was present in each of the four classrooms to conduct the experiment, along with one other researcher.

Design

The experimental design for Main Study II was a 3 x 4 randomized block design with three levels of reading achievement (high, medium, and low) and four treatment groups.

Previous reading achievement was based on scores on the Reading section (Test R) of the Iowa Tests of Basic Skills which was administered in October 1971. By blocking on previous level of reading

Table 10

Content of Lessons

for Main Study II

Concept	Version	Content
Population	Form S-101	4 + instances* and 3 - instances* plus concept definition
	Form S-102	4 + instances* plus concept definition
	Form S-103	2 + instances plus concept definition
Habitat	Form S-201	4 + instances* and 2 - instances* plus concept definition
	Form S-202	4 + instances* plus concept definition
	Form S-203	2 + instances plus concept definition
Community	Form S-301	4 + instances* and 3 - instances* plus concept definition
	Form S-302	4 + instances* plus concept definition
	Form S-303	2 + instances plus concept definition

* Indicates the number of instances which is recommended by Markle and Tiemann.

achievement, an equal number of Ss were randomly assigned to each of the four treatment conditions. A summary of the treatment conditions in Main Study II is provided in Table 11, and the number of subjects in each cell is given in Table 12. Independent variables in this experiment were number of concept instances presented and type of concept instances presented (in the presence of a concept definition).

Table 11

Summary of Treatment Conditions

for Main Study II

Treatment	Independent Variable			
	Number of Instances		Definition	Emphasis
	Positive	Negative		
Cond. 1	Markle-Tiemann*	Markle-Tiemann	yes	no
Cond. 2	Markle-Tiemann	none	yes	no
Cond. 3	2	none	yes	no
Cond. 4	Placebo Lessons			

*According to the Markle-Tiemann paradigm, the specific number of positive and negative instances depends upon the number of relevant and irrelevant attributes which a concept possesses.

Table 12

Number of Subjects By Cell

for Main Study II

Previous Reading Achievement Level*	Treatment Condition				Totals Across Condition
	MT**positives MT negatives plus concept definition	MT positives plus concept definition	2 positives plus concept definition	Control placebo lessons	
High	10	10	10	10	40
Medium	10	10	10	8	38
Low	10	8	10	8	36
Totals Across Reading Level	30	28	30	26	114

* Based on reading scores (Test R) of the Iowa Tests of Basic Skills

** Refers to the rational set of instances as recommended by Markle and Tiemann

Hypotheses

The following specific hypotheses were made for Main Study II (refer to Table 11 for a summary of treatment conditions):

Ia. On the dependent variable "recognition of new concept instances", Conditions 1, 2, and 3 will not be significantly different from each other, but they will result in significantly better perfromance than Condition 4 (Control).

b. On the dependent variable "overgeneralization", Condition 1 will not be significantly different from Condition 2.

II. On the dependent variable "knowledge of concept definition", Conditions 1, 2, and 3 will not be significantly different from each other, but they will result in significantly better performance than Condition 4 (Control).

III. On the dependent variable "knowledge of interrelationships among concepts", Conditions 1, 2, and 3 will not be significantly different from each other, but they will result in significantly better performance than Condition 4 (Control).

MAIN STUDY III

The purpose of the third main study was to investigate the effects of definition and emphasis of relevant (defining) attributes on the acquisition of environmental concepts.

Subjects

Subjects for Main Study III were 117 sixth-grade students from four different schools located in the same joint school district in which Main Study II was conducted. The final number of subjects used in analysis was 114, due to absences at the time of retention testing. Ss for studies II and III comprised the entire sixth-grade population of seven elementary schools in the joint district. Since the student population of one school was involved in both studies II and III, the children were randomly divided into two equal groups, so that one half of the sixth-grade population participated in study II and the other half participated in study III. The Ss were unfamiliar with the concepts in the experimental lessons, and according to the teachers there had been no formal instruction on the concepts population, habitat, or community.

Materials

Lessons

The three concept lessons for this experiment were variations of the same lessons which were used in the Pilot Study (with modifications as described in Chapter 3). A collection of the envirnomental lessons including concept definitions and emphasis of relevant (defining) attributes in a modified linear programming format comprises Appendix D. Concept definitions for population, habitat, and community were given, and emphasis of relevant attributes was provided in a modified linear programming format with immediate feedback. Lesson content for the treatment groups in Main Study III is shown in Table 13. Placebo lessons were the Conservation, Transportation, and Invention lessons which were used in the previous two main studies.

Tests

Test 1 (Populations), Test 2 (Habitats), Test 3 (Communities), and Test 4 (Final Test) were the same tests which were administered in Main Studies I and II. The long-term Retention Test was composed of Tests 1-4 combined into a single booklet.

Procedure

The procedure for this study was the same as the procedure which was outlined for Main Study I. It was necessary for the proctor to briefly explain the use of cardboard strips to conceal feedback in the margins of the programmed booklets. Consult Table 6 for the schedule of lessons and tests for Main Study III.

Two proctors (E and another researcher familiar with the study) were present in each of the four classrooms.

Table 13
Content of Lessons
for Main Study III

Concept	Version	Content
Population	Form S-101	4 + instances* and 3 - instances*
	Form S-102	4 + instances* and 3 - instances* with concept definition
	Form S-103	4 + instances* and 3 - instances* with concept definition and emphasis of relevant attributes**
Habitat	Form S-201	4 + instances* and 2 - instances*
	Form S-202	4 + instances* and 2 - instances* with concept definition
	Form S-203	4 + instances* and 2 - instances* with concept definition and emphasis of relevant attributes**
Community	Form S-301	4 + instances* and 3 - instances*
	Form S-302	4 + instances* and 3 - instances* with concept definition
	Form S-303	4 + instances* and 3 - instances* with concept definition and emphasis of relevant attributes**

* Indicates the number of instances which is recommended by Markle and Tiemann.

** Indicates use of modified linear programming format.

Design

The experimental design for Main Study III was a 3 x 4 randomized block design with three levels of reading achievement (high, medium, and low) and four treatment groups.

Previous reading achievement was based on scores from the Iowa Tests of Basic Skills (Test R: Reading), administered in October 1971. Blocking on previous reading achievement resulted in a random assignment of an equal number of Ss to each of the four treatment conditions. A summary of the treatment conditions for Main Study III is given in Table 14, and the number of subjects in each cell is given in Table 15. The only independent variables manipulated in this study were concept definition and emphasis of relevant attributes, so the number of positive and negative instances remained constant across treatments.

Table 14
Summary of Treatment Conditions
for Main Study III

Treatment	Independent Variable			
	Number of Instances		Definition	Emphasis
	Positive	Negative		
Cond. 1	Markle-Tiemann*	Markle-Tiemann	no	no
Cond. 2	Markle-Tiemann	Markle-Tiemann	yes	no
Cond. 3	Markle-Tiemann	Markle-Tiemann	yes	yes
Cond. 4	Placebo lessons			

* According to the Markle-Tiemann paradigm, the specific number of positive and negative instances depends upon the number of relevant and irrelevant attributes which a concept possesses.

Table 15

Number of Subjects By Cell

for Main Study III

Previous Reading Achievement Level*	Treatment Condition: MT**positives MT negatives	MT positives MT negatives plus concept definition	MT positives MT negatives plus definition and emphasis	Control placebo lessons	Totals Across Condition
High	8	8	9	9	34
Medium	9	9	9	8	35
Low	7	9	8	9	33
Totals Across Reading Level	24	26	26	26	102

* Based on reading scores (Test R) of the Iowa Tests of Basic Skills

** Refers to the rational set of instances as recommended by Markle and Tiemann

Hypotheses

The following specific hypotheses were made for Main Study III (refer to Table 14 for a summary of treatment conditions):

I. On the dependent variable "recognition of new concept instances", Condition 3 will result in significantly better performance than Condition 2, which will be significantly better than Condition 1, which will be significantly better than Condition 4 (Control).

II. On the dependent variable "knowledge of concept definition", Condition 3 will result in significantly better performance than Condition 2, which will be significantly better than Condition 1, which will be significantly better than Condition 4 (Control).

III. On the dependent variable "knowledge of interrelationships among concepts", Condition 3 will result in significantly better performance than Condition 2, which will be significantly better than Condition 1, which will be significantly better than Condition 4 (Control).

RESULTS

Scores on each of the four immediate acquisition tests (Test 1, Test 2, Test 3, Test 4) and the retention test were obtained for each S. Analysis of the data was performed using the following dependent variables: recognition of new concept instances (examples and non-examples), overgeneralization, undergeneralization, knowledge of concept definition, and knowledge of interrelationships among concepts. Since recognition of new instances was tested in Part 1 of Tests 1, 2, and 3, the raw scores on each test were converted to Z-scores and the dependent measure used in analysis was the sum of Z-scores across all three tests. Overgeneralization and undergeneralization were based on errors made on the recognition items, so these dependent measures were also sums of Z-scores across Tests 1, 2, and 3. Knowledge of concept definition was determined from one multiple-choice item on Test 1, one multiple-choice item on Test 2, one multiple-choice item on Test 3, and a three-part matching question on Test 4. For this reason, the dependent measure for knowledge of concept definition was determined in the following manner: (a) raw scores for the Test 1 item, the Test 2 item, and the Test 3 item were summed and then converted to a Z-score; (b) raw scores for the three items in Test 4 were summed and converted to a Z-score; (c) the dependent measure used in analysis was a sum of the two Z-scores. Knowledge of interrelationships among concepts was tested with 5 true-false items and 4 multiple-choice itmes contained in Test 4. Since all items were contained in a single test and there was no need to combine scores across tests, the dependent measure for knowledge of interrelationships was left as a raw score total. For reference, raw score data (means and standard deviations by test) for all dependent variables in the three main studies are provided in Appendix H.

Prior to testing specific ordered hypotheses for the main studies, a 3 x 4 analysis of variance (3 levels of previous reading achievement and 4 treatment conditions) was performed on each of the five dependent variables (recognition of new concept instances, over-generalization, undergeneralization, knowledge of definition, and knowledge of interrelationships). A complete set of ANOVA tables for all three main studies (immediate acquisition and retention) comprises Appendix F.

Where specific hypotheses were made concerning the three dependent variables recognition of new instances, knowledge of definition, and knowledge of interrelationships, non-orthogonal planned comparisons were carried out with a selected overall alpha (α) of .05 so that each of the three pairwise contrasts was tested with α = .016. The Tukey procedure for pairwise comparisons of mean scores was used for the overgeneralization and undergeneralization variables (employing a simultaneous error rate of .05). As cell sizes were unequal, an approximate critical value based on the harmonic mean was used.

For organizational purposes, the results of each main study to be reported in this chapter will be presented by dependent variable.

In general, the effect of stratifying on reading achievement was significant for every dependent variable on immediate acquisition ($p<.030$) as well as retention ($p<$ [illegible]). There were no significant stratification x treatment interactions in Main Study I for either immediate acquisition or retention ($p>.631$). Means for each of the dependent variables in Main Study I are given by treatment condition and stratification level in immediate acquisition and retention Tables (Appendix G).

Psychometric Characteristics of the Tests

On Main Study I immediate acquisition, the Hoyt reliability estimate for the 60 items (Tests 1-3) testing recognition of new concept instances was quite high (.81). Reliabilities for the 6 items testing knowledge of definition and the 9 items testing knowledge of interrelationships were much lower (.20 and .46 respectively).

Recognition of New Instances

According to Markle and Tiemann as well as Merrill, both positive and negative instances are important in teaching concepts because positive instances promote generalization to new examples and negative instances promote discrimination of new non-examples. Any deviation from the prescribed rational set of instances, such as exclusion of negative instances from lesson materials, or reduction of the number of positive instances, would be expected to significantly depress performance. In Main Study I the first treatment condition received a full set of both positive and negative instances, the second condition received the prescribed number of positive instances without negative instances, the third condition received only 2 positive instances (or half the prescribed number) without negative instances, and the fourth condition was the control group which received "placebo" lessons. It was hypothesized that on this dependent variable (recognition of new concept instances) Condition 1 would result in significantly better performance than Condition 2, which would be significantly better than Condition 3, which in turn would be significantly better than Condition 4 (Control).

Based on a selected overall α of .05 (each of the three directional pairwise contrasts with an α of .016), it was found that on immediate acquisition Condition 1 with the Markle-Tiemann set of positive and negative instances resulted in significantly better recognition than Condition 2 without negative instances ($t=2.6$, $p<.005$), and the Condition 3 with only two positive instances resulted in significantly better recognition than Condition 4, Control ($t=2.6$, $p<.005$). The difference in recognition between Condition 2 (full number of positive instances, no negatives) and Condition 3 (only 2 positive instances, no negatives) was significant at the .05 level on immediate acquisition but not in the predicted direction, indicating that the larger number of positive instances without negative instances may have been confusing. On the retention measure given 16 days following immediate acquisition, Condition 1 was still found to be significantly better than Condition 2 ($t=2.3$, $p<.01$), but Condition 3 [illegible] Condition 4 [illegible], $p < .10$). Means for [illegible] not [illegible] the predicted direction ($t = 1.$[illegible]).

An alternate [illegible] the Tukey procedure with [illegible] .05 [illegible] six pairwise comparisons (critical [illegible] immediate acquisition, and 1.70 for retention) [illegible] immediate acquisition Condition 1 resulted [illegible] new instances than Condition 2 (difference [illegible]), Condition 1 was better than Condition 4, [illegible] 1.97), and Condition 3 was also better than Control ([illegible] 1.45). On retention the only significant pairwise contrast was between Condition 1 and Control (difference of 1.86).

Overgeneralization

Because of the function of negative instances to promote discrimination, it would be expected that the exclusion of negative instances would result in overgeneralization (S does not discriminate and identifies nonexamples as examples). It was hypothesized that Condition 2, which lacked negative instances, would result in significantly more overgeneralization than Condition 1 with both positive and negative instances. Using the Tukey procedure with a simultaneous error rate of .05 for six pairwise contrasts (critical values = 1.60 for immediate acquisition, and 1.81 for retention), it was found that Condition 2 did in fact result in significantly more overgeneralization than Condition 1. Condition 1 with the full set of positive and negative instances was clearly superior to all other conditions (1 different from 2 by 2.54; 1 different from 3 by 1.79; 1 different from 4 by 1.73). There were no other pairwise contrasts significantly different at $\alpha = .05$ on immediate acquisition. On retention, Condition 1 was still significantly different from Condition 2 (by 2.02), indicating that removal of negative instances caused stable overgeneralization but none of the other pairwise comparisons was significant.

Undergeneralization

No specific hypothesis was made for the dependent variable of undergeneralization because this classification error is said to occur only when a full number of negative instances is presented with an incomplete set of positive instances. Tukey comparisons were made using a simultaneous error rate of .05 for the six pairwise contrasts (critical values = 1.45 for immediate acquisition, and 1.67 for retention). Condition 3 with only two positive instances (no negatives) was found to result in significantly less undergeneralization than either Condition 1 (by 1.49) or Condition 4, Control (by 2.14) on immediate acquisition. On the retention test none of the six pairwise comparison was significant.

Knowledge of Definition

As one of the components of formal concept attainment, it was expected that knowledge of the concept definition (defining attributes) would be significantly better for conditions utilizing the Markle-Tiemann number of positive and negative instances than the conditions lacking the rational set of instances. It was hypothesized that on the dependent variable knowledge of concept definition (each of three pairwise contrasts tested with alpha of .016) Condition 1 would result in significantly better performance than Condition 2, Condition 2 would be significantly better than Condition 3, and Condition 3 would be significantly better than Condition 4 (Control).

None of the three directional pairwise contrasts was significant on immediate acquisition or retention ($p > .10$). The means for Conditions 3 and 4 were not found to be in the predicted direction, since Control Ss did better than Condition 3 Ss on both immediate acquisition and retention items testing knowledge of definition.

Using the Tukey procedure with a simultaneous error rate of .05 for six pairwise comparisons, none of the contrasts was significant on immediate acquisition or retention.

Knowledge of Interrelationships

It was hypothesized that Condition 1 would result in significantly better knowledge of interrelationships than Condition 2, which would be significantly better than Condition 3, which in turn would be significantly better than Condition 4 (Control). None of the three directional pairwise contrasts was significant on immediate acquisition or retention ($p > .10$). Ss in Condition 2 did perform slightly better than Ss in Condition 1 on knowledge of interrelationships ($t=-.17$ on retention as well as immediate acquisition). For this reason the means for Conditions 1 and 2 were not in the predicted direction.

No significant differences were obtained on immediate acquisition or retention when Tukey comparisons were made (using a simultaneous error rate of .05).

MAIN STUDY II

In the second main study the effect of stratifying on reading achievement was highly significant at the .0001 level for every dependent variable on immediate acquisition, except for knowledge of interrelationships which was significant at the .0[illegible] level. There were two significant stratification X treatment interactions on immediate acquisition for recognition of new instances ($p<.064$) and undergeneralization ($p<.048$). On the retention measure stratification on reading achievement was highly significant for recognition of new instances, overgeneralization, and undergeneralization ($p<.0001$), while knowledge of definition was significant at the .0005 level. The effect of stratifying on reading achievement was not significant on the retention test for knowledge of interrelationships ($p<.239$), and on this dependent variable there was also a significant interaction ($p<.030$) of stratification with treatments. One other stratification X treatment interaction was significant on retention for recognition of new instance. ($p<.058$).

Though [illegible] the present research, unusually [illegible] low [illegible] Ss in Main Study II may partially account for the stratification X treatment interactions which [illegible] each of the dependent variables in Main Study II are given by treatment condition and stratification level in immediate acquisition and retention Tables (Appendix C).

Psychometric Characteristics of the Tests

On Main Study II immediate acquisition, the Hoyt reliability estimate for the 60 items (Tests 1-3) which tested recognition of new examples and non-examples was very high (.88). The reliability of the [illegible] items which tested knowledge of definition was somewhat lower (.53), while items testing knowledge of interrelationships (5 true-false questions and 4 multiple-choice questions contained in Test 4) were found to be quite unreliable (.18).

Recognition of New Instances

Because of the amount of information contained in a concept definition, it was expected that when a definition was provided there would be no significant differences among conditions due to the number of positive and negative instances presented. All four treatment conditions in this study were identical to the conditions in Main Study I except for the inclusion of a definition (statement of

[illegible]

[illegible] new instances [illegible] the first [illegible]

result in [illegible] Condition 4 (Control).

[illegible] retention, a de- [illegible] each of the fol- lowing [illegible]

[illegible]

On immediate [illegible] recognition of new instances [illegible] to be significantly better than [illegible] the three conditions were not found to be [illegible] each other ($F_{2,102}$ [illegible]) [illegible] On the retention [illegible] immediate acquisition Condition [illegible] significantly better than Control [illegible] conditions were also significantly [illegible] ($F_{2,102}$ [illegible], p [illegible] .01). Therefore, pairwise [illegible] 1, 2, and 3 were made using the Tukey [illegible] rate of .05. The pairwise comparison [illegible] that Condition 1 with the full set of positive and negative [illegible] definition was significantly worse than [illegible] the negative instances (the difference in means [illegible] the established critical value of 1.32). No other pairwise differences among Conditions 1, 2, and 3 were detected.

Overgeneralization

It was hypothesized that in Main Study II when a definition was included there would be no significant difference in overgeneralization between Condition 2 (lacking negative instances) and Condition 1 (with both positive and negative instances). Using the Tukey procedure with a simultaneous error rate of .05 for six pairwise comparisons (critical values = 1.33 for immediate acquisition, and 1.41 for retention), it was determined that Condition 2 was not significantly different from Condition 1 on overgeneralization. Conditions 1, 2, and 3 were all found to result in significantly less overgeneralization than the control group (1 different from 4 by 1.77; 2 different from 4 by 2.14; 3 different from 4 by 1.41), but on retention only Conditions 2 and 3 resulted in significantly less overgeneralization than the control group (2 different from 4 by 2.13; 3 different from 4 by 1.41).

Undergeneralization

Since there were no treatment conditions specifically designed to result in the classification error of undergeneralization, no hypothesis was made for this dependent variable. Tukey comparisons among the four treatment conditions (simultaneous error rate of .05; critical values = 1.24 for immediate acquisiton, and 1.36 for retention) revealed no significant pairwise differences on immediate acquisition, but on retention Condition 1 was found to result in significantly more undergeneralization than Condition 2 (by 1.38). This pairwise difference, [illegible] in words [illegible] indicates that [illegible] and negative [illegible] so much when asked [illegible] the retention test.

Knowledge of [illegible]

[illegible] conditions 1, 2, and [illegible] would be no significant difference [illegible] on knowledge of definition. Since Condition [illegible] only "placebo" lessons without the concept definition [illegible] Conditions 1, 2, and 3 were hypothesized to result in [illegible] better performance on knowledge of definition than Condition [illegible] (control).

On immediate [illegible] Conditions 1, 2, and 3 were found to result in significant[illegible] knowledge of definition than Condition 4 (t=[illegible].86, p<.005), but the three conditions were also significantly different [illegible] each other (F_{102}=[illegible] p<.005). Therefore, pairwise contrasts among Conditions 1, 2, and 3 were performed using the Tukey procedure with a simultaneous error rate of .05 (critical value for immediate [illegible] .91). It was found that Condition 1 with the full set of positive and negative instances (plus definition) resulted in significantly poorer knowledge of definition than Condition 2 which lacked the negative instances (difference in means was 1.2[illegible]). None of the other pairwise differences was significant on immediate acquisition. Although failing to meet the selected alpha (α) of .025 which was used to test the 1 = 2 = 3 > 4 hypothesis, Conditions 1, 2, and 3 were found to be significantly better than Condition 4 on the retention measure at the .05 level (t=1.72).

An alternate analysis of the data was performed using the Tukey procedure with a simultaneous error rate of .05 for six pairwise contrasts (critical values = .86 for immediate acquisition, and .98 for retention). On immediate acquisition it was found that Condition 1 was significantly worse than Condition 2 (difference of −1.21), and that Conditions 2 and 3 were significantly better than Control (differences of 1.37 and .82 respectively). None of the six pairwise comparisons was significant on retention.

Knowledge of Interrelationships

It was hypothesized that on this dependent variable, Conditions 1, 2, and 3 would not differ significantly from each other while resulting in better knowledge of interrelationships than Condition 4 (Control). On immediate acquisition no significant differences were found to exist among Conditions 1, 2, and 3 ($F_{2,102}$=1.20, p >.10), but neither were these three conditions significantly different from the control group, Condition 4 (t=−1.13, p>.10). On the retention test the treatments were still not different from each other ($F_{2,102}$=.267, p>.10) and not different from control (t=−.71, p>.10).

No significant differences were obtained on immediate acquisition or retention when Tukey comparisons were made (using a simultaneous error rate of .05).

MAIN STUDY III

In the third main study the effect of stratifying on reading achievement was significant at the .0001 level for recognition of new instances, undergeneralization, and knowledge of definition on both immediate acquisition and retention. Stratification effects for overgeneralization and knowledge of interrelationships were less highly significant ($p < .003$ on immediate acquisition, and $p < .0002$ on retention). With the exception of knowledge of definition ($p < .056$) and knowledge of interrelationships ($p < .088$), there were no significant stratification X treatment interactions in Main Study III on immediate acquisition. There were no significant interactions for any of the dependent variables on the retention measure administered two weeks following immediate acquisition.

This experiment was designed to investigate the effects of definition and emphasis of relevant attributes upon the five dependent variables when a full set of positive and negative instances (as recommended by Markle and Tiemann) w[illegible]i. Means for each of the dependent variables in the t[illegible] study are given by treatment condition and stratification level in immediate acquisition and retention Tables (Appendix G).

[illegible]

On Main Study (1) immediate acquisition the [illegible] reliability estimate for [illegible] to identify [illegible] concept instances (2[illegible] items in [illegible], [illegible] items in Test 2, and 1[illegible] items in Test 3) was .8[illegible] [illegible] reliabilities for the 6 items testing knowledge of definition and the 9 items testing knowledge of interrelationships were .66 and .54 respectively.

Recognition of New Instances

Since Condition [illegible] was the most complete condition in terms of informational content, it was hypothesized that Condition 3 would result in significantly better recognition than Condition 2, which would result in significantly better recognition than Condition 1, which in turn would be significantly better than Condition 4 (Control). Based on a selected overall α of .05 (each of the three directional pairwise contrasts with α=.016), it was found that Condition 3 (exemplars/non-exemplars plus concept definition and emphasis of relevant attributes) resulted in better recognition than Condition 2 (exemplars/non-exemplars plus concept definition, without emphasis); even though the difference was not significant at the seclected .016 level, Condition 3 was better than Condition 2 at $p<.025$ (t=2.02). On the retention measure this same relationship was significant at the .0[illegible] level (t=2.35). On both immediate acquisition and retention measures the differences between Condition

2 and Condition 1 (exemplars/non-exemplars only) were found to be significant, but not in the predicted direction. Ss in Condition 2 who received instances plus definition actually performed worse than Ss in Condition 1 who received instances only ($t=-3.40$, $p < .0005$ on immediate acquisition; $t=-2.17$, $p < .025$ on retention). On both immediate acquisition and retention, Condition 1 resulted in significantly better recognition performance than Control ($t=4.95$, $p < .0005$ on immediate acquisition; $t=3.42$, $p < .0005$ on retention).

An alternate analysis of the data using the Tukey procedure with a simultaneous error rate of .05 (critical values = 1.38 for immediate acquisition, and 1.47 for retention) revealed that on immediate acquisition Conditions 1, 2, and 3 were each found to be significantly better than Control (differences of 3.21, 1.39, and 2.45 respectively). It was also found that Condition 1 (with exemplars/non-exemplars only) resulted in significantly better recognition of new instances than Condition 2 (with exemplars/non-exemplars plus definition). On retention the only contrasts still significant were between Condition 1 and Control, and between Condition 3 and Control.

Overgeneraliztion

No specific hypothesis concerning overgeneralization was made for the third main study because all treatment conditions except control included the full rational set of positive and negative instances, which according to Markle and Tiemann does not result in overgeneralization errors. Using the Tukey procedure with a simultaneous error rate of .05 for the six pairwise contrasts (critical value = 1.53 for immediate acquisition, and 1.59 for retention) it was found that on immediate acquisition Condition 2 (exemplars/non-exemplars plus definition) resulted in significantly more overgeneraliztion than either Condition 3 (difference of 1.98) or Condition 1 (difference of 3.18). Additionally, Condition 1 was found to result in significantly less overgeneralization than Control (difference of 2.57) on immediate acquisition. On the retention test given fourteen days after immediate acquisition, all of the relationships described for immediate acquisition were still significant, and Condition 3 was also found to result in significantly less overgeneralization than Control (difference of 1.61).

Undergeneralization

As with overgeneralization, no specific hypothesis for this dependent variable was made. Tukey pairwise comparisons were performed using a simultaneous error rate of .05 for the six contrasts (critical values = 1.30 for immediate acquisition, and 1.61 for retention). On immediate acquisition, Conditions 1, 2, and 3 were each found to result in significantly less undergeneralization than Condition 4, Control (Condition 1 different by 1.66; Condition 2 different by 2.81; Condition 3 different by 2.13). On retention, Condition 2 was found to result in less undergeneralization than Condition 4 (difference of 1.70), and Condition 2 resulted in less undergeneralization than Condition 1 (difference of 1.53). None of the other pairwise differences was significant.

Knowledge of Definition

It was hypothesized that Condition 3 would result in significantly better knowledge of concept definition than Condition 2, which would be significantly better than Condition 1, which in turn would be significantly better than Condition 4 (Control). Only one of these three directional pairwise contrasts was found to be significant on immediate acquisition at the selected alpha (α) of .016: Condition 1 was significantly better than Control (t=2.58, p <.001). On retention the same relationship was significant at the .10 level (t=1.49). Ss in condition 2 (which included a definition of the concept) surprisingly performed worse on recognition of concept definition than Ss in Condition 1 (with only exemplars/non-exemplars). For this reason the means for Conditions 1 and 2 were not in the predicted direction (t=-.41, p >.10 for immediate acquisition; t= -.08, p >.10 for retention).

An alternate analysis of the data was performed using Tukey comparisons (with a simultaneous error rate of .05 for six pairwise contrasts; critical values = .96 for immediate acquisition, and 1.03 for retention). On immediate acquisition, Conditions 1, 2, and 3 each resulted in better recognition of the concept definition than Control (differences of 1.01, .98, and 1.47 respectively). On retention none of the six pairwise contrasts was significant.

Knowledge of Interrelationships

On knowledge of interrelationships it was hypothesized that

Condition 3 would be significantly better than Condition 2, which would be significantly better than Condition 1, which would be significantly better than Control, Condition 4. None of the three directional pairwise contrasts was significant on immediate acquisition or retention ($p > .10$).

No significant differences were obtained on immediate acquisition or retention when Tukey comparisons were made (using a simultaneous error rate of .05).

VI

CONCLUSIONS AND DISCUSSION

CONCLUSIONS

Three main studies were carried out to determine the effects of certain instructional variables (number of positive and negative teaching instances, presentation of a concept definition which specified the relevant attributes, and emphasis of the relevant or defining attributes) on the attainment of three environmental concepts at the formal level.

The findings of the three main studies were in general:

Main Study I

Presenting the number of teaching examples and non-examples recommended by Markle and Tiemann was found to result in significantly better recognition of new instances than presentations which did not include non-examples. Removal of negative instances from the lessons caused significantly greater overgeneralization (exactly as demonstrated by Tennyson, 1971).

Main Study II

When a definition which specified the defining attributes of the concept was presented, the number of teaching instances did not significantly affect performance (as found by Frayer, 1970). At the time of retention testing, Ss who had read lessons containing a Markle-Tiemann number of exemplars/non-exemplars as well as definition made significantly more errors of undergeneralization than Ss who had read lessons which were alike except for the removal of negative instances.

Main Study III

Subjects who received a rational set of exemplars/non-exemplars plus definition performed less well than subjects who received only exemplars/non-exemplars or subjects who received exemplars/non-exemplars plus definition and emphasis of relevant attributes. In general, presenting a concept definition did not facilitate performance, and providing emphasis of relevant attributes was effective only as contrasted with a conditon which included the Markle-Tiemann set of exemplars/non-exemplars plus concept definition.

DISCUSSION

Main Study I

The purpose of the first main study was to determine the effects of number of positive and negative instances on the acquisiton of formal-level environmental concepts in the absence of a definition which specified the relevant or defining attributes of the concept.

One of the dependent variables of interest was recognition of new instances, as measured by the Ss's ability to identify unencountered instances of the concept as examples or non-examples. Treatment Condition 1 included a rational set of both positive and negative instances as recommended by Markle and Tiemann (1969): there were four positive instances to cover the range of irrelevant attributes, and three negative instances (similar to exemplars in irrelevant attributes) to exclude each of the three relevant attributes of the concept to be taught. Condition 2 was identical to Condition 1 except for the removal of the three negative instances, and Condition 3 included only two positive instances without negative instances, while Condition 4 was Control. According to Markle and Tiemann as well as Merrill (1971), both examples and non-examples are important in concept attainment, so Condition 1 was hypothesized to constitute the best condition. With the exception of the dependent variables undergeneralization and knowledge of interrelationships, Condition 1 means across the three reading levels were consistently better than the means for the other three conditions. Of great importance was the finding that on immediate acquisition as well as retention the presentation of a full set of positive and negative instances resulted in significantly better recognition of new instances than a presentation in which the negative instances had been removed, just as Markle and Tiemann (1969) and Merrill (1971) hypothesized and Tennyson (1971) empirically demonstrated. It was found that when negative instances were not included in the teaching sequence, the different number of positive instances (four in Condition 2, and

2 in Condition 3) did not significantly affect recognition of new examples and non-examples. While Condition 1 with the full set of instances and non-instances differed significantly from Control on both immediate acquisition and retention, the other two treatment conditions lacking negative instances generally resulted in performance not significantly different from Control. For this reason it would appear that additional positive instances above two related to the present concepts is not important when a complement of negative instances is not also included.

An interesting trend in the data obtained for recognition of new instances on both immediate acquisition and retention was the ordering of Condition 3 (with only two positive instances and no negatives) better than Condition 2 (with the full Markle-Tiemann number of positives and no negatives). Since the full rational set consisted of four positive instances, it was hypothesized that Condition 2 with the full Markle-Tiemann number of positives would result in significantly better recognition of new instances than Condition 3 with only two positives, or half the Markle-Tiemann number. Obtained scores on this variable for the two conditions, however, were not in the predicted direction. Although the difference in means was statistically significant only on immediate acquisition ($p < .05$), the poor performance of Condition 2 Ss seems to demonstrate that unless positive instances are complemented with negative instances (as recommended by Markle and Tiemann), an increase in the number of instances only increases the informational load without strengthening discrimination, resulting in confusion on the part of the student. This is also suggested by the fact that Ss who received the Markle-Tiemann set of positive instances without negatives actually overgeneralized more than any other treatment group, including Control.

Another important aspect of the data analysis was the pattern of classification errors which emerged. Markle and Tiemann postulate that failure to provide an adequate set of negative instances will result in errors of overgeneralization, in which S does not properly discriminate many of the unencountered non-examples from examples. In the present study it was preciesly this error of overgeneralization which occurred when negative examples were removed: Condition 2, which was identical to Condition 1 except for the removal of negative instances, resulted in significantly more overgeneralization ($p < .05$) on immediate acquisition as well as retention (administered 16 days later). Additionally, undergeneralization was not found to be significantly different between Conditions 1 and 2, which indicates that removal of the negative instances led specifically to overgeneralization and not random responding in general. The finding in the present study that removal of negative instances resulted in significantly poorer recognition of new instances and significantly more overgeneralization errors is in agreement with the findings of Tennyson (1971). Evidently non-examples facilitate concept learning performance by causing students to focus on the relevant attributes of the concept, as Tennyson suggested.

On immediate acquisition, it was discovered that presenting the full set of positive and negative instances (Condition 1) resulted in significantly more undergeneralization than presenting just two positive instances (Condition 2) and that the control group (Condition 4) also resulted in more errors of undergeneralization than Condition 2. No significant pairwise differences were detected at the time of retention testing, however. The fact that undergeneralization was higher for Condition 1 (with the Markle-Tiemann number of positive and negative instances) than for any other condition except Control is not really unreasonable, in view of what Merrill and Tennyson have written. They postulate that the more information presented, the more conservative the responses: Ss are more willing to reject a true instance than accept a false one (or more inclined to undergeneralize slightly than to overgeneralize). The results of this study provide substantial empirical support for the findings of Tennyson (1971) concerning the value of negative instances as well as for the theoretical formulations of Markle and Tiemann concerning the role of examples and non-examples in teaching concepts effectively.

In the first main study no significant effects were found for treatment conditions on either of the two remaining dependent measures, knowledge of concept definition or knowledge of interrelationships. This lack of significance may have been partly due to the difficult nature of the concepts themselves and to the fairly large number of relevant and irrelevant attributes associated with each of the three environmental concepts (population, habitat, and

and community). The low reliability of the test items (.20 for knowledge of definition and .46 for knowledge of interrelationships) was probably a critical factor which influenced results on the dependent measures dealing with definition and relationships among concepts.

Main Study II

The second main study was essentially a replication of Main Study I with the addition of a concept definition (which specified the defining attributes) to treatment conditions 1, 2, and 3. The purpose of the study was largely to find out whether the minimum number of teaching examples and non-examples presented might be reduced when a concept definition is provided. Intuitively, the additional information which a definition supplies concerning relevant attributes of the concept should encourage the student to focus upon the relevant attributes, thus reducing the demands for other information-providing variables, specifically the number of positive and negative instances.

It was found that on immediate acquisition, recognition of new concept instances for Conditions 1, 2, and 3 was significantly better than Control ($p < .001$), but the three treatment conditions were not significantly different from each other ($p > .10$). For this reason, defining a concept by specifying its relevant attributes may actually decrease the number of positive and negative instances required. On retention, only Condition 2 (with the Markle-Tiemann set of positives, no negatives) and Condition 3 (two positives, no negatives) were found to be significantly better on recognition of new concept instances than Control. Performance on this dependent variable was surprisingly better on retention ($p < .05$) for Condition 2, without negative instances, than for Condition 1 which included the Markle-Tiemann number of negative as well as positive instances. At first glance, this long-term effect seems rather inconsistent with the predictions of Markle and Tiemann, who maintain that negative instances have a decided facilitative effect on concept attainment. A closer look at the pattern of errors, however, reveals that Ss receiving the full set of examples and non-examples also undergeneralized significantly more than Ss who did not receive non-examples. The finding that the presentation of a rational set of positive instances alone resulted in better recognition than the presentation of a rational set of both positive and negative instances may actually be a function of the amount of information which was supplied during instruction. As pointed out in the discussion of Main Study I, Merrill and Tennyson have found that the more information presented, the more cautious the responses tend to be, or the more reluctant the Ss are to identify low-probability examples. In terms of patterns of classification errors, an overly cautious pattern of responses is known as undergeneralization, and it was precisely this tendency to undergeneralize which resulted in worse performance by Condition 1 Ss than Condition 2 Ss (who received less information) on long-term recognition of new concept instances.

A complete analysis of the pattern of errors made in Main Study II showed that on immediate acquisition Condition 1 (with the full Markle-Tiemann set of positive and negative instances plus definition) resulted in significantly more undergeneralization than Condition 3 (with only two positive instances plus definition), and that Condition 4 (Control) resulted in significantly more errors of undergeneralization than Condition 3. Only the relationship between Conditions 1 and 3 was still significant on retention, in addition to the significantly greater undergeneralization for Condition 1 as contrasted with Condition 2 which was discussed in the previous paragraph.

The pattern of overgeneralization errors was somewhat more consistent across immediate acquisition, with Control resulting in significantly more overgeneralization than any of the other three treatment conditions. A most interesting finding was that removal of negative instances (Condition 2) did not lead to more overgeneralization than presentation of the full Markle-Tiemann set of positive and negative instances (Condition 1). According to Markle and Tiemann, of course, the removal of negative instances should result in greater overgeneralization because the student fails to discriminate properly if he has not encountered a minimum of non-examples during instruction. Tennyson (1971) proposed that negative instances cause the student to focus on the relevant or defining attributes of the concept, and in his study the removal of negative instances caused significantly more overgeneralization even when a definition of the concept <u>adverb</u> was included in all treatments except control. Based on results of the present study, there is evidence that the learning of three particular environmental concepts may be different when a definition of the concept is provided than when the definition is omitted. In the case where negative instances were not included in the teaching sequence, presenting a definition seemed to compensate for the lack of non-examples. Since a definition states the defining or relevant attributes of the concept it may be that when the concepts to be taught are somewhat difficult, merely presenting positive instances along with the concept definition is as effective as presenting the full set of positive and negative instances without definition.

Tennyson found significant differences among treatments even when a definition of the concept was included. There is evidence that one reason for the unusual results obtained in the present study may be the difficulty level of the instances which were used for teaching and for testing. While Tennyson used a range of instances during instruction, the present study presented teaching examples and non-examples which were nearly equivalent in terms of difficulty (approximately at the .70 probability level). From an examination of those items which were missed by students in Main Study II who overgeneralized or undergeneralized significantly, it was found that the items most generally misclassified were low-probability items (ranging from probabilities of .23 to .57). Performance on the dependent measures (which included a range of instances) might have been somewhat different if the full range of examples and non-examples had been presented in the lessons.

For the dependent variable knowledge of definition on immediate acquisition it was found that Condition 1 Ss (who received both positive and negative instances plus definition) actually performed worse ($p < .05$) than Condition 2 Ss (who received only positive instances plus definition). This may have been a result of the amount of information presented, and the group receiving the definition plus positive as well as negative instances could have been more confused by the great amount of information than the group receiving less information. On immediate acquisition there was also significantly better knowledge of definition for Condition 2 (Markle-Tiemann positives plus definition) and for Condition 3 (two positives plus definition) than for Control. There were no significant differences obtained for any pairwise comparisons on the retention measure administered eleven days following immediate acquisition. The reliability of the six items which tested knowledge of definition was not as high as the reliability of items measuring recognition of new instances (.53 and .88 respectively) yet it was sufficiently high to be acceptable.

No significant differences were obtained for knowledge of interrelationships on either immediate acquisition or retention. As in Main Study I, this was probably due to the extremely low reliability of test items which were designed to measure knowledge of interrelationships among concepts (for Main Study II the Hoyt reliability estimate was .18 for the nine items testing knowledge of interrelationships).

It should be noted that in Main Study II on both immediate acquisition and retention there was a very extreme spread between the scores obtained by high reading achievement Ss and low reading achievement Ss. Tables in Appendix G show that on nearly every dependent variable the low stratification subjects performed poorly. Thus the importance of reading and vocabulary on performance is clearly demonstrated.

Main Study III

The third main study focused on the effects of presenting the Markle-Tiemann set of positive and negative instances as contrasted with presenting a Markle-Tiemann number of instances with concept definition, and with a concept definition plus emphasis of the relevant (defining) attributes.

On immediate acquisition as well as retention there was a significant difference between Condition 3 (which presented exemplars/non-exemplars as well as definition and emphasis of relevant attributes of the concepts population, habitat, and community significantly facilitated the learning of these three environmental concepts ($p < .025$ on immediate acquisition; $p < .01$ on retention). The condition which included instances plus definition and emphasis was not found to be significantly different from the condition which included only the rational set of examples and non-examples (Condition 1), however. On immediate acquisition, children in each of Conditions 1, 2, and 3 performed significantly better than

Control on recognition of new instances, and while Condition 2 Ss did not differ significantly from Control on retention, the children in Condition 1 who received the rational set of instances and in Condition 3 who received the rational set of instances plus definition and emphasis were still found to perform better ($p < .05$) than the control group on the retention measure given 14 days after immediate acquisition.

Quite unexpectedly, the condition which included only the rational set of examples and non-examples (Condition 1) was found to result in better recognition of new instances than the condition which included the examples and non-examples as well as a concept definition (Condition 2). This was true on immediate acquisition ($p < .0005$) as well as retention ($p < .025$). A close look at the classification errors revealed that the poor performance of Condition 2 Ss was obviously due to a greater incidence of overgeneralization. As shown in Appendix G, students in Condition 2 who received exemplars/non-exemplars plus definition made more overgeneralization errors than any other treatment group, including Control. Since undergeneralization was not also greater, the inferior recognition of new concept instances must be solely attributable to failure to discriminate new non-examples from examples. This finding is not consistent with Main Study II in which the presentation of a Markle-Tiemann set of positive and negative instances plus definition resulted in correct classification rather than overgeneralization. Neither is this finding consistent with the recommendations of Markle and Tiemann. It is quite possible that the content or wording of the definitions themselves may have confused the students in Condition 2, whereas students in Condition 1 did not receive the definition and students in Condition 3 received clarification of the relevant attributes (emphasis) in a format which provided immediate feedback to correct false assumptions. Although it is difficult to determine why Condition 2 Ss may have been confused, some aspect of the definition must have complicated recognition of new concept instances, because even students in the control group who presumably responded on a random basis made fewer overgeneralization errors than students who received the exemplars/ non-exemplars plus definition.

Aside from the finding that Condition 2 resulted in significantly more overgeneralization than Condition 1, it was found that Ss in Condition 1 (who received only the rational set of examples and non-examples) made significantly fewer overgeneralization errors than the Control Ss on both immediate acquisition and retention. Presenting exemplars/non-exemplars with definition and emphasis of relevant attributes (Condition 3) resulted in fewer errors of overgeneralization than presenting exemplars/non-exemplars with definition in the absence of emphasis (Condition 2) on both immediate acquisition and retention ($p < .05$). Additionally, Condition 3 was found to result in significantly less overgeneralization than Control on the retention measure.

Analysis of the patterns of classification errors also revealed that Ss in Conditions 1, 2, and 3 undergeneralized less than the control subjects on immediate acquisition, but on retention the only difference which was still significant existed between Condition 2 and Control. Subjects who received only exemplars/non-exemplars were also found to make fewer errors of undergeneralization than subjects who were apparently confused by the presentation of exemplars/non-exemplars plus concept definition.

Generally, emphasis of relevant attributes was found to be a facilitative technique only in contrast to Condition 2 in which Ss were possibly confused and performed very poorly. Because the significant pairwise differences between Condition 2 and other conditions may have been due to weaknesses in the experimental treatments, only limited conclusions should be drawn. The specific definitions of environmental concepts used in this study were not found to facilitate concept learning performance on either immediate acquisition or retention.

The only significant differences found for knowledge of definition were on immediate acquisition: Conditions 1, 2, and 3 each resulted in better performance than Control. Another indication of the possible confusion exhibited by Ss who received exemplars/non-exemplars plus definition was the fact that Condition 1 (which included only exemplars/non-exemplars) actually resulted in slightly better recognition of definition than Condition 2 which clearly specified the concept definition. (A Hoyt reliability of .66 was obtained for items testing knowledge of definition in Main Study III.)

Performance on the remaining dependent variable, knowledge of interrelationships, was not found to be significantly different for any pairwise comparisons between treatments. Possible weaknesses in in the instrument (item reliability of .54) as well as difficulty of the concepts and complexity of the interrelationships involved were probably factors which influenced performance on this particular variable.

Limitations of the Studies

It must be pointed out that the specific concepts which were chosen for the present research are critical factors which necessarily limit the generalizability of results. Wherever concept definition was studied as a dependent variable, the particular concept definition as stated in the lesson booklets was also a major determinant of the final experimental outcomes. Associated with each of the three environmental concepts population, habitat, and community were approximately 3 relevant attributes and 4 irrelevant attributes. If other concepts of different levels of difficulty and different numbers of relevant and irrelevant attributes had been chosen, the results of the present study might have been quite different. It is important to reiterate that the number of relevant and irrelevant attributes of a concept (a) determines the number of positive and negative instances used in teaching, (b) affects the length and amount of information in the definition, and (c) regulates the frequency of emphasis during instruction. In turn, the number of relevant and irrelevant attributes influences the results obtained when number of instances, definition, and emphasis are the independent variables under investigation. It is likely that the difficulty of the teaching and testing instances and emphasis of irrelevant as well as relevant attributes might also be important independent variables in concept attainment, although the present study was not designed to test either of these.

The degree to which the actual experimental meterials influenced results must also be considered. Hoyt reliability estimates for the items testing recognition of new instances were quite high for all three studies (.81, .88, and .88 respectively), but items measuring knowledge of definition and knowledge of interrelationships were much less reliable.

In summary, based on results of the present study it seems that complexity of the concept to be taught and the number of relevant and irrelevant attributes of the concept have a powerful influence upon concept attainment, especially when the concepts are drawn from typical school subject-matters.

REFERENCES

Amster, H. Effect of instructional set and variety of instances on children's learning. Journal of Educational Psychology, 1966, 57, 74-85.

Anderson, R. C. & Kulhavy, R. W. Learning concepts from definitions. American Educational Research Journal, 1972, in press.

Blount, N. S., Klausmeier, H. J., Johnson, S. L., Fredrick, W. C., & Ramsay, J. G. The effectiveness of programmed materials in English syntax and the relationship of selected variables to the learning of concepts. Technical Report Number 17. Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, Madison, 1967.

Bourne, L. E., Ekstrand, B. R., & Dominowski, R. L. The psychology of thinking. Englewood Cliffs, N. J.: Prentice-Hall, 1971.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. A study of thinking. New York: John Wiley & Sons, 1956.

Callantine, M. F. & Warren, J. M. Learning sets in human concept formation. Psychological Reports, 1955, 1, 363-367.

Clark, D. Cecil. Teaching concepts in the classroom: A set of teaching prescriptions derived from experimental research. Journal of Educational Psychology, 1971, 62, 253-278.

Frayer, D. A. Effects of number of instances and emphasis of relevant attribute values on mastery of geometric concepts by fourth- and sixth-grade children. Technical Report Number 116. Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, Madison, 1970.

Fryatt, M. & Tulving, E. Interproblem transfer in identification of concepts involving positive and negative instances. Canadian Journal of Psychology, 1963, 17, 106-117.

Gagné, R. M. The conditions of learning. New York: Holt, Rinehart and Winston, 1965.

Hovland, C. I. & Weiss, W. Transmission of information concerning concepts through positive and negative instances. Journal of Experimental Psychology, 1953, 45, 175-182.

Huttenlocher, J. Some effects of negative instances on the formation of simple concepts. Psychological Reports, 1962, 11, 35-42.

Huttenlocher, J. Development of formal reasoning on concept formation problems. Child Development, 1964, 35, 1233-1242.

Klausmeier, H. J. Cognitive operations in concept learning. Paper presented at the meeting of the American Psychological Association, Washington, D. C., September, 1971.

Klausmeier, H. J., Davis, J. K., Ramsay, J. G., Fredrick, W. C., & Davies, M. H. Concept learning and problem solving: A bibliography, 1950-1964. Technical Report Number 1. Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, Madison, 1965.

Markle, S. M. & Tiemann, P. W. Really understanding concepts. Champaign, Illinois: Stipes, 1969.

Masilela, A. Effects of positive and negative instances on the transfer of verbal concepts. Unpublished doctoral dissertation, University of California, Los Angeles, 1964.

Mechner, F. Science education and behavioral technology. In R. Glaser (Ed.), Teaching machines and programmed learning, II: Data and direction. Washington, D. C.: National Education Association, 1965, 441-507..

Merrill, M. D. Necessary psychological conditions for defining instructional outcomes. In M. D. Merrill (Ed.), Instructional design readings. Englewood Cliffs, N. J.: Prentice Hall, 1971.

Merrill, M. D., & Tennyson, R. D. Concept acquisition as a function of relationships between positive and negative instances. Final Report of USOE Small Contract O-H-014, September, 1971.

Morrisett, L. Jr. & Hovland, C. I. A comparison of three varieties of training in human problem solving. Journal of Experimental Psychology, 1959, 58, 52-55.

Olson, L. Concept attainment of high school sophomores. Journal of Educational Psychology, 1963, 54, 213-216.

Remstad, R. C. Optimizing the response to a concept attainment test through sequential classroom experimentation. Technical Report Number 85. Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, Madison, 1969.

Scott, J. A. The effects of short- and long-term retention and on transfer of two methods of presenting selected geometry concepts. Technical Report Number 138. Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, Madison, 1970.

Smoke, K. L. Negative instances in concept learning. Journal of Experimental Psychology, 1933, 16, 583-588.

Smuckler, N. S. Concept formation as a function of method of presentation and ratio of positive to negative instances. Technical Report Number 27. Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, Madison, 1967.

Tennyson, R. D. Correct classification, overgeneralization, undergeneralization, and misconception as a function of probability, divergency, and matching of instances in concept acquisition. Working Paper Number 19. Division of Communication Services, Brigham Young University, Provo, Utah, 1971.

Tennyson, R. D., Woolley, F. R., & Merrill, M. D. Exemplar and non-exemplar variables which produce correct classification behavior and specified classification errors. Journal of Educational Psychology, 1972, 63, 144-152.

Appendices B through D have been omitted from this publication, but are available on microfilm from Memorial Library, University of Wisconsin, Madison, Wisconsin

APPENDIX A

GENERAL INSTRUCTIONS TO STUDENTS

Instructions to Students

Good afternoon.

My name is ____________. I am working with some people at the University of Wisconsin in Madison. We're very interested in finding better ways to help children learn about science. Today you will be able to help us by answering some questions. You will be taking a short test, but there is no reason to worry about the test; just do your best, and you will help us to make learning about science easier for other boys and girls.

I am going to hand out the booklets now. Please don't open your booklet until I tell you to do so. [Distribute booklets.]

Now write your name, and the name of your teacher on the cover, and fill in today's date. [write the date on the blackboard] Now turn to the next page where it says Word List. These are some of the words which you may find in the questions you are going to answer. Because some of them are a little unusual, let's take just a minute to go over them. Do you see the word ______ ? What number is it ? [Good!] Let's pronounce it together. What about the word ______ ? Do you see it listed here ? [Continue until all of the words have been pronounced, and their numbers indicated.] There may be other words in the questions which look new to you, but don't be afraid to raise your hand if you need help. I will come around and pronounce any words which are giving you trouble.

Let's turn to the next page now. This is just for practice [the Sample Page] but the rest of the test looks like this, so it is important to learn how to mark your answers. [Read through the instructions, and allow the children to mark sample items, then show on the blackboard how their answers should look.]

Unless anyone has a question for me to answer now, I think that you are ready to start on the test. Any questions ?

When you finish, just keep your booklets and work quietly on something else until everyone is through. O.K., you may turn the page and begin.

APPENDIX F

SUMMARY TABLES FOR UNIVARIATE ANALYSES OF VARIANCE PERFORMED ON THE FIVE DEPENDENT VARIABLES IN MAIN STUDIES I, II, AND III

Main Study I

Immediate Acquisition

Variable	Source	df	MS	F	p
Recognition of New Concept Instances	Reading Level	2	43.74	11.60	.0001
	Treatment	3	19.69	5.22	.002
	Interaction (R x T)	6	1.55	.41	.871
	Error (Within cells)	85	3.77		
	Total	96			
Overgeneralization	Reading Level	2	16.42	3.65	.030
	Treatment	3	26.88	5.98	.001
	Interaction (R x T)	6	1.65	.37	.898
	Error (Within cells)	85	4.49		
	Total	96			
Undergeneralization	Reading Level	2	32.16	8.67	.0004
	Treatment	3	21.56	5.82	.001
	Interaction (R x T)	6	2.69	.72	.631
	Error (Within cells)	85	3.71		
	Total	96			
Knowledge of Definition	Reading Level	2	28.15	15.09	.0001
	Treatment	3	1.75	.94	.426
	Interaction (R x T)	6	.39	.21	.973
	Error (Within cells)	85	1.87		
	Total	96			
Knowledge of Interrelationships	Reading Level	2	29.24	11.64	.0001
	Treatment	3	.19	.08	.973
	Interaction (R x T)	6	1.72	.69	.662
	Error (Within cells)	85	2.51		
	Total	96			

Main Study I

Retention

Variable	Source	df	MS	F	p<
Recognition of New Concept Instances	Reading Level	2	49.92	9.78	.0002
	Treatment	3	9.97	1.95	.127
	Interaction (R x T)	6	2.14	.42	.865
	Error (Within cells)	85	5.10		
	Total	96			
Overgeneral-ization	Reading Level	2	21.62	3.73	.028
	Treatment	3	16.30	2.81	.044
	Interaction (R x T)	6	2.97	.51	.797
	Error (Within cells)	85	5.80		
	Total	96			
Undergeneral-ization	Reading Level	2	23.62	4.81	.011
	Treatment	3	7.66	1.56	.206
	Interaction (R x T)	6	1.92	.39	.883
	Error (Within cells)	85	4.91		
	Total	96			
Knowledge of Definition	Reading Level	2	22.30	9.94	.0002
	Treatment	3	.97	.43	.729
	Interaction (R x T)	6	1.07	.48	.823
	Error (Within cells)	85	2.24		
	Total	96			
Knowledge of Interrelation-ships	Reading Level	2	18.34	9.07	.0003
	Treatment	3	1.36	.67	.572
	Interaction (R x T)	6	1.29	.64	.701
	Error (Within cells)	85	2.02		
	Total	96			

Main Study II

Immediate Acquisition

Variable	Source	df	MS	F	p<
Recognition of New Concept Instances	Reading Level	2	120.68	34.91	.0001
	Treatment	3	28.62	8.28	.0001
	Interaction (R x T)	6	7.13	2.06	.064
	Error (Within cells)	102	3.46		
	Total	113			
Overgeneralization	Reading Level	2	74.79	20.51	.0001
	Treatment	3	24.35	6.68	.0004
	Interaction (R x T)	6	5.43	1.49	.190
	Error (Within cells)	102	3.65		
	Total	113			
Undergeneralization	Reading Level	2	43.23	13.69	.0001
	Treatment	3	9.76	3.09	.031
	Interaction (R x T)	6	6.99	2.21	.048
	Error (Within cells)	102	3.16		
	Total	113			
Knowledge of Definition	Reading Level	2	34.62	22.53	.0001
	Treatment	3	10.57	6.88	.0003
	Interaction (R x T)	6	1.09	.71	.644
	Error (Within cells)	102	1.54		
	Total	113			
Knowledge of Interrelationships	Reading Level	2	5.47	3.35	.039
	Treatment	3	1.99	1.22	.306
	Interaction (R x T)	6	1.65	1.01	.421
	Error (Within cells)	102	1.63		
	Total	113			

Main Study II

Retention

Variable	Source	df	MS	F	p<
Recognition of New Concept Instances	Reading Level	2	137.11	41.91	.0001
	Treatment	3	25.24	7.72	.0002
	Interaction (R x T)	6	6.91	2.11	.058
	Error (Within cells)	102	3.27		
	Total	113			
Overgeneral-ization	Reading Level	2	81.76	19.78	.0001
	Treatment	3	21.49	5.20	.002
	Interaction (R x T)	6	4.97	1.20	.311
	Error (Within cells)	102	4.13		
	Total	113			
Undergeneral-ization	Reading Level	2	59.96	17.42	.0001
	Treatment	3	11.10	3.22	.026
	Interaction (R x T)	6	5.90	1.71	.125
	Error (Within cells)	102	3.44		
	Total	113			
Knowledge of Definition	Reading Level	2	16.54	8.28	.0005
	Treatment	3	6.97	3.49	.018
	Interaction (R x T)	6	3.61	1.81	.110
	Error (Within cells)	102	2.00		
	Total	113			
Knowledge of Interrelation-ships	Reading Level	2	3.83	1.45	.239
	Treatment	3	.92	.35	.789
	Interaction (R x T)	6	6.47	2.45	.030
	Error (Within cells)	102	2.64		
	Total	113			

Main Study III

Immediate Acquisition

Variable	Source	df	MS	F	p<
Recognition of New Concept Instances	Reading Level	2	63.54	18.00	.0001
	Treatment	3	47.11	13.34	.0001
	Interaction (R x T)	6	5.88	1.67	.139
	Error (Within cells)	90	3.53		
	Total	101			
Overgeneralization	Reading Level	2	26.58	6.14	.003
	Treatment	3	48.49	11.20	.0001
	Interaction (R x T)	6	4.91	1.13	.349
	Error (Within cells)	90	4.33		
	Total	101			
Undergeneralization	Reading Level	2	46.34	14.81	.0001
	Treatment	3	37.91	12.12	.0001
	Interaction (R x T)	6	3.07	.98	.442
	Error (Within cells)	90	3.13		
	Total	101			
Knowledge of Definition	Reading Level	2	32.75	17.05	.0001
	Treatment	3	9.74	5.07	.003
	Interaction (R x T)	6	4.12	2.14	.056
	Error (Within cells)	90	1.92		
	Total	101			
Knowledge of Interrelationships	Reading Level	2	22.47	7.37	.001
	Treatment	3	.59	.19	.901
	Interaction (R x T)	6	5.81	1.91	.088
	Error (Within cells)	90	3.05		
	Total	101			

Main Study III

Retention

Variable	Source	df	MS	F	p<
Recognition of New Concept Instances	Reading Level	2	89.02	22.32	.0001
	Treatment	3	22.54	5.65	.001
	Interaction (R x T)	6	4.05	1.02	.420
	Error (Within cells)	90	3.99		
	Total	101			
Overgeneralization	Reading Level	2	44.54	9.51	.0002
	Treatment	3	38.35	8.19	.0001
	Interaction (R x T)	6	3.95	.84	.541
	Error (Within cells)	90	4.68		
	Total	101			
Undergeneralization	Reading Level	2	52.74	11.04	.0001
	Treatment	3	17.28	3.62	.016
	Interaction (R x T)	6	4.10	.86	.529
	Error (Within cells)	90	4.78		
	Total	101			
Knowledge of Definition	Reading Level	2	44.72	21.85	.0001
	Treatment	3	1.95	.95	.419
	Interaction (R x T)	6	3.39	1.66	.141
	Error (Within cells)	90	2.05		
	Total	101			
Knowledge of Interrelationships	Reading Level	2	20.44	9.95	.0002
	Treatment	3	.45	.22	.883
	Interaction (R x T)	6	2.29	1.12	.359
	Error (Within cells)	90	2.05		
	Total	101			

APPENDIX G

MEANS FOR DEPENDENT VARIABLES BY TREATMENTS AND PREVIOUS READING ACHIEVEMENT LEVELS ON IMMEDIATE ACQUISITION AND RETENTION FOR STUDIES I, II, AND III

Main Study I, Immediate Acquisition: Means* for Dependent Variables by treatments (three combinations of instances) and Previous Reading Achievement Levels

Dependent Variable	Previous Reading Achievement Level	Treatment Condition			
		1 (N=24) MT**positives MT negatives	2 (N=24) MT positives only	3 (N=26) two positives only	4 (N=23) Control
Recognition of New Instances (MSE= 3.77)	High (N=30)	2.63	.38	1.51	.63
	Medium (N=35)	.71	- .56	.07	- .99
	Low (N=32)	- .32	-1.51	- .02	-2.54
	Across Reading Level (N=97)	1.01	- .60	.48	- .97
Overgeneralization (MSE= 4.50)	High (N=30)	-2.80	.49	- .71	- .19
	Medium (N=35)	-1.07	1.11	.75	- .18
	Low (N=32)	- .68	1.38	.67	1.11
	Across Reading Level (N=97)	-1.52	1.02	.27	.21
Undergeneralization (MSE= 3.71)	High (N=30)	- .67	-1.54	-1.50	- .75
	Medium (N=35)	.45	- .52	-1.04	1.81
	Low (N=32)	1.50	1.03	- .70	1.96
	Across Reading Level (N=97)	.42	- .30	-1.06	1.07
Knowledge of Definition (MSE= 1.87)	High (N=30)	1.50	1.05	.34	.68
	Medium (N=35)	.36	.20	- .07	.11
	Low (N=32)	- .74	-1.06	-1.02	-1.19
	Across Reading Level (N=97)	.37	.03	- .27	- .11
Knowledge of Interrelationships (MSE= 2.51)	High (N=30)	6.63	6.86	6.50	6.00
	Medium (N=35)	6.00	6.44	7.00	6.44
	Low (N=32)	5.25	4.86	4.33	5.00
	Across Reading Level (N=97)	5.96	6.04	5.93	5.87

*Means for the first four dependent variables are given as combined Z-scores; means for the last dependent variable are given as raw scores.

**Refers to the number of instances recommended by Markle and Tiemann.

Main Study I, Retention: Means* for Dependent Variables by Treatments (three combinations of instances) and Previous Reading Achievement Levels

Dependent Variable	Previous Reading Achievement Level	Treatment Condition 1 (N=24) MT** positives MTnegatives	2 (N=24) MT positives only	3 (N=26) 2 positives only	4 (N=23) Control
Recognition of New Instances (MSE=5.10)	High (N=30)	2.80	.75	1.05	1.15
	Medium (N=35)	.28	- .84	- .34	- .32
	Low (N=32)	- .39	-1.57	- .31	-2.05
	Across Reading Level (N=97)	.90	- .62	.10	- .40
Overgeneralization (MSE=5.80)	High (N=30)	-2.28	.07	- .52	-1.21
	Medium (N=35)	- .45	1.11	1.05	- .53
	Low (N=32)	- .28	1.75	- .01	1.08
	Across Reading Level (N=97)	-1.01	1.02	.20	- .24
Undergeneralization (MSE=4.91)	High (N=30)	-1.43	-1.06	-1.00	- .31
	Medium (N=35)	.26	.02	- .94	1.32
	Low (N=32)	1.24	.31	.31	1.17
	Across Reading Level (N=97)	.02	- .20	- .53	.78
Knowledge of Definition (MSE=2.24)	High (N=30)	1.14	.72	.68	1.09
	Medium (N=35)	- .50	.30	- .33	.26
	Low (N=32)	- .27	- .70	-1.15	-1.03
	Across Reading Level (N=97)	.12	.09	- .30	.12
Knowledge of Interrelationships (MSE=2.02)	High (N=30)	7.13	7.29	7.25	6.43
	Medium (N=35)	6.25	6.67	6.44	6.67
	Low (N=32)	6.13	5.88	5.11	5.00
	Across Reading Level (N=97)	6.50	6.58	6.23	6.09

*Means for the first four dependent variables are given as combined Z-scores; means for the last dependent variable are given as raw scores.

**Refers to the number of instances recommended by Markle and Tiemann.

Main Study II, Immediate Acquisition: Means* for Dependent Variables by Treatments (three combinations of instances plus concept definition) and Previous Reading Achievement Levels

Dependent Variable	Previous Reading Achievement Level	1 (N=30) MT**positives MT negatives & definition	2 (N=28) MT positives & definition	3 (N=30) two positives & definition	4 (N=26) Control
Recognition of New Instances (MSE=3.46)	High (N=40)	2.15	1.74	2.87	- .01
	Medium (N=38)	.69	.08	.57	-1.83
	Low (N=36)	-2.84	.07	-1.52	-2.88
	Across Reading Level (N=114)	- .01	.67	.64	-1.46
Overgeneralization (MSE=3.65)	High (N=40)	-2.25	-1.22	-1.81	.28
	Medium (N=38)	-1.04	- .98	.38	1.41
	Low (N=36)	2.07	.02	1.29	2.66
	Across Reading Level (N=114)	- .41	- .78	- .05	1.36
Undergeneralization (MSE=3.16)	High (N=40)	- .85	-1.24	-2.13	- .31
	Medium (N=38)	.17	1.62	- .93	.63
	Low (N=36)	1.86	- .24	.71	1.00
	Across Reading Level (N=114)	.39	.07	- .79	.38
Knowledge of Definition (MSE=1.54)	High (N=40)	.29	1.45	1.09	- .04
	Medium (N=38)	- .02	.64	.85	- .21
	Low (N=36)	-1.58	.10	-1.27	-1.67
	Across Reading Level (N=114)	- .43	.76	.23	- .60
Knowledge of Interrelationships (MSE=1.63)	High (N=40)	6.70	6.60	6.50	7.00
	Medium (N=38)	6.60	6.50	6.30	7.13
	Low (N=36)	5.60	7.00	5.60	5.88
	Across Reading Level (N=114)	6.30	6.68	6.13	6.69

* Means for the first four dependent variables are given as combined Z-scores; means for the last dependent variable are given as raw scores.

** Refers to the number of instances recommended by Markle and Tiemann.

Main Study II, Rentention: Means* for Dependent Variables by Treatments (three combinations of instances plus concept definition) and Previous Reading Achievement Levels

Dependent Variable	Previous Reading Achievement Level	Treatment Condition			
		1 (N=30) MT**positives MT negatives & definition	2 (N=28) MT positives & definition	3 (N=30) two positives & definition	4 (N=26) Control
Recognition of New Instances (MSE=3.27)	High (N=40)	1.68	2.39	3.13	.62
	Medium (N=38)	- .25	.01	.13	-1.60
	Low (N=36)	-2.85	.68	-1.96	-2.69
	Across Reading Level (N=114)	- .47	1.05	.43	-1.08
Overgeneralization (MSE=4.13)	High (N=40)	-1.59	-1.53	-2.31	.19
	Medium (N=38)	- .34	- .64	- .07	.66
	Low (N=36)	1.81	- .45	1.85	3.09
	Across Reading Level (N=114)	- .04	- .91	- .18	1.23
Undergeneralization (MSE=3.44)	High (N=40)	- .73	-1.80	-1.79	-1.25
	Medium (N=38)	.98	1.07	.27	1.42
	Low (N=36)	2.44	- .80	.46	- .17
	Across Reading Level (N=114)	.90	- .49	- .35	- .10
Knowledge of Definition (MSE=2.00)	High (N=40)	- .09	.77	1.39	.28
	Medium (N=38)	- .08	- .04	.97	- .70
	Low (N=36)	-1.19	.55	-1.08	-1.00
	Across Reading Level (N=114)	- .46	.42	.43	- .41
Knowledge of Interrelationships (MSE=2.64)	High (N=40)	5.40	6.50	6.60	7.10
	Medium (N=38)	6.70	5.50	6.60	5.63
	Low (N=36)	5.50	6.75	5.00	6.00
	Across Reading Level (N=114)	5.87	6.21	6.07	6.31

* Means for the first four dependent variables are given as combined Z-scores; means for the last dependent variable are given as raw scores.

** Refers to the number of instances recommended by Markle and Tiemann

Main Study III, Immediate Acquisition: Means* for Dependent Variables by Treatments (Markle-Tiemann set of positive and negative instances alone, with definition, or with definition and emphasis of relevant attributes) and Previous Reading Achievement Levels

Dependent Variable	Previous Reading Achievement Level	Treatment Condition			
		1 (N=24) MT**instances alone	2 (N=26) MT instances & definition	3 (N=26) MT instances, definition, & emphasis	4 (N=26) Control
Recognition of New Instances (MSE=3.53)	High (N=34)	3.85	.09	2.61	- .56
	Medium (N=35)	.66	.20	.49	-2.31
	Low (N=33)	- .20	-1.26	-1.16	-2.40
	Across Reading Level (N=102)	1.47	- .34	.71	-1.73
Overgeneralization (MSE=4.33)	High (N=34)	-3.67	1.23	-1.87	.46
	Medium (N=35)	- .92	.98	- .52	.68
	Low (N=33)	- .70	1.98	.84	1.25
	Across Reading Level (N=102)	-1.77	1.41	- .57	.80
Undergeneralization (MSE=3.13)	High (N=34)	-1.53	-1.66	-1.91	.23
	Medium (N=35)	.23	-1.40	- .62	2.45
	Low (N=33)	1.41	- .48	1.30	2.35
	Across Reading Level (N=102)	- .01	-1.16	- .48	1.65
Knowledge of Definition (MSE=1.92)	High (N=34)	1.91	.54	1.62	.18
	Medium (N=35)	- .67	.78	.73	-1.63
	Low (N=33)	- .80	- .92	- .67	-1.23
	Across Reading Level (N=102)	.15	.12	.61	- .86
Knowledge of Interrelationships (MSE=3.05)	High (N=34)	6.75	5.75	6.78	6.56
	Medium (N=35)	4.78	7.00	6.33	5.25
	Low (N=33)	5.29	4.56	4.50	5.11
	Across Reading Level (N=102)	5.58	5.77	5.92	5.65

* Means for the first four dependent variables are given as combined Z-scores; means for the last dependent variable are given as raw scores.

** Refers to the number of positive and negative instances recommended by Markle and Tiemann

Main Study III, Retention: Means* for Dependent Variables by Treatments (Markle-Tiemann set of positive and negative instances alone, with definition, or with definition and emphasis of relevant attributes) and Previous Reading Achievement Levels

Dependent Variable	Previous Reading Achievement Level	Treatment Condition: 1 (N=24) MT**instances alone	2 (N=26) MT instances & definition	3 (N=26) MT instances, definition, & emphasis	4 (N=26) Control
Recognition of New Instances (MSE=3.99)	High (N=34)	3.33	.74	2.53	.28
	Medium (N=35)	- .28	- .38	1.20	-1.21
	Low (N=33)	- .77	-1.53	-1.37	-2.53
	Across Reading Level (N=102)	.78	- .43	.87	-1.15
Overgeneralization (MSE=4.68)	High (N=34)	-3.32	.53	-1.67	- .28
	Medium (N=35)	- .94	1.62	-1.51	.81
	Low (N=33)	.30	1.68	.89	1.82
	Across Reading Level (N=102)	-1.37	1.31	- .83	.78
Undergeneralization (MSE=4.78)	High (N=34)	-1.23	-1.88	-2.10	- .18
	Medium (N=35)	1.77	-1.39	- .02	.77
	Low (N=33)	1.09	.29	1.38	1.63
	Across Reading Level (N=102)	.57	- .96	- .31	.74
Knowledge of Definition (MSE=2.05)	High (N=34)	1.75	.59	1.94	.51
	Medium (N=35)	- .32	.03	.29	- .50
	Low (N=33)	- .92	- .48	-1.84	-1.24
	Across Reading Level (N=102)	.19	.03	.20	- .41
Knowledge of Interrelationships (MSE=2.05)	High (N=34)	6.88	6.50	7.44	7.22
	Medium (N=35)	5.67	6.00	6.33	6.38
	Low (N=33)	5.57	6.11	4.75	5.44
	Across Reading Level (N=102)	6.04	6.19	6.23	6.35

* Means for the first four dependent variables are given as combined Z-scores; means for the last dependent variable are given as raw scores.

** Refers to the number of positive and negative instances recommended by Markle and Tiemann.

APPENDIX H

RAW SCORE MEANS AND STANDARD DEVIATIONS FOR DEPENDENT MEASURES ON IMMEDIATE ACQUISITION AND RETENTION BY TREATMENT CONDITION WITHIN READING ACHIEVEMENT LEVEL

Raw Score Means and Standard Deviations* for Main Study I (Immediate Acquisition)

Test	Dependent Variable	Previous Reading Achievement Level	Treatment Condition 1 (N=24)	2 (N=24)	3 (N=26)	4 (Control) (N=23)	Totals Across Condition
Test 1, Part A	Recognition of New Instances and Non-Instances of the Concept Population	High (N=30)	18.25 (2.63)	15.29 (2.55)	17.25 (5.07)	16.14 (2.17)	16.80 (3.54)
		Medium (N=35)	15.50 (2.74)	16.78 (2.04)	17.33 (2.79)	14.56 (2.41)	16.06 (2.74)
		Low (N=32)	16.13 (2.71)	15.75 (2.17)	16.56 (1.64)	11.43 (2.44)	15.13 (3.00)
		Across Reading Level (N=97)	16.63 (2.94)	16.00 (2.33)	17.04 (3.41)	14.09 (3.01)	
	Overgeneralization for the Concept Population	High (N=30)	3.00 (1.73)	8.00 (2.33)	5.88 (2.37)	6.14 (2.10)	5.67 (2.80)
		Medium (N=35)	6.00 (3.04)	6.11 (2.18)	5.89 (2.51)	5.44 (2.06)	5.86 (2.47)
		Low (N=32)	5.13 (2.32)	6.63 (3.12)	6.56 (1.17)	6.57 (1.99)	6.22 (2.33)
		Across Reading Level (N=97)	4.17 (2.73)	6.83 (2.69)	6.12 (2.12)	6.00 (2.11)	
	Undergeneralization for the Concept Population	High (N=30)	3.75 (2.17)	0.43 (0.73)	1.88 (3.48)	2.71 (2.60)	2.23 (2.77)
		Medium (N=35)	3.38 (3.20)	2.11 (1.59)	1.78 (2.20)	5.00 (2.94)	3.06 (2.85)

next page . . .

*Note: Standard deviations are given in parentheses.

		Low (N=32)	3.75 (2.17)	2.63 (2.45)	1.78 (1.40)	6.57 (2.67)	3.53 (2.81)
		Across Reading Level (N=97)	3.63 (2.56)	1.79 (1.98)	1.81 (2.47)	4.78 (3.15)	
Test 1, Part B	Knowledge of Definition of Population	High (N=30)	0.50 (0.50)	0.14 (0.35)	0.13 (0.33)	0.43 (0.49)	0.30 (0.46)
		Medium (N=35)	0.25 (0.43)	0.22 (0.42)	0.00 (0.00)	0.11 (0.31)	0.14 (0.35)
		Low (N=32)	0.13 (0.33)	0.38 (0.48)	0.33 (0.47)	0.00 (0.00)	0.22 (0.41)
		Across Reading Level (N=97)	0.29 (0.45)	0.25 (0.43)	0.15 (0.36)	0.17 (0.38)	
Test 2, Part A	Recognition of New Instances and Non-Instances of the Concept Habitat	High (N=30)	19.50 (4.80)	15.71 (3.84)	16.63 (4.92)	17.71 (4.49)	17.43 (4.77)
		Medium (N=35)	18.38 (4.58)	12.56 (1.50)	13.44 (2.31)	13.56 (4.95)	14.37 (4.24)
		Low (N-32)	15.38 (4.27)	13.13 (2.26)	15.33 (3.62)	12.71 (4.59)	14.22 (3.95)
		Across Reading Level (N=97)	17.75 (4.88)	13.67 (2.94)	15.08 (3.94)	14.57 (5.16)	
	Overgeneralization for the Concept Habitat	High (N=30)	4.50 (5.10)	8.00 (3.70)	7.75 (5.24)	6.29 (4.62)	6.60 (4.94)
		Medium (N=35)	4.25 (4.18)	11.00 (2.58)	10.67 (2.91)	8.56 (3.92)	8.74 (4.33)
		Low (N=32)	6.13 (3.22)	9.25 (2.28)	8.89 (3.38)	9.71 (3.49)	8.47 (3.42)

next page . . .

		Across Reading Level (N=97)	4.96 (4.32)	9.54 (3.12)	9.15 (4.09)	8.22 (4.25)	
	Undergeneralization for the Concept Habitat	High (N=30)	1.00 (1.22)	1.29 (1.67)	0.63 (0.48)	0.71 (0.70)	0.90 (1.14)
		Medium (N=35)	2.25 (2.90)	1.33 (1.33)	0.89 (1.29)	2.89 (2.42)	1.83 (2.22)
		Low (N=32)	3.50 (2.24)	2.63 (2.60)	0.78 (1.23)	2.29 (2.86)	2.25 (2.49)
		Across Reading Level (N=97)	2.25 (2.45)	1.75 (2.03)	0.77 (1.09)	2.04 (2.40)	
Test 2, Part B	Knowledge of Definition of Habitat	High (N=30)	0.25 (0.43)	0.57 (0.49)	0.38 (0.48)	0.57 (0.49)	0.43 (0.50)
		Medium (N=35)	0.75 (0.43)	0.11 (0.31)	0.33 (0.47)	0.56 (0.50)	0.43 (0.49)
		Low (N=32)	0.13 (0.33)	0.38 (0.48)	0.11 (0.31)	0.29 (0.45)	0.22 (0.41)
		Across Reading Level (N=97)	0.38 (0.48)	0.33 (0.47)	0.27 (0.44)	0.48 (0.50)	
Test 3, Part A	Recognition of New Instances and Non-Instances of the Concept Community	High (N=30)	8.63 (1.11)	7.57 (1.84)	8.25 (1.71)	6.57 (1.76)	7.80 (1.80)
		Medium (N=35)	6.88 (2.62)	6.00 (2.16)	6.56 (1.64)	6.11 (2.51)	6.37 (2.28)
		Low (N=32)	5.63 (1.22)	4.38 (0.99)	6.00 (1.41)	5.29 (2.05)	5.34 (1.57)
		Across Reading Level (N=97)	7.04 (2.17)	5.92 (2.16)	6.88 (1.85)	6.00 (2.23)	

next page . . .

	Overgeneralization for the Concept Community	High (N=30)	1.13 (1.27)	2.00 (1.69)	1.50 (1.66)	2.71 (1.28)	1.80 (1.60)
		Medium (N=35)	2.13 (1.69)	3.11 (1.29)	2.78 (1.47)	2.33 (1.25)	2.60 (1.48)
		Low (N=32)	2.63 (1.58)	3.88 (0.93)	2.89 (1.20)	3.29 (1.98)	3.16 (1.52)
		Across Reading Level (N=97)	1.96 (1.65)	3.04 (1.51)	2.42 (1.57)	2.74 (1.57)	
	Undergeneralization for the Concept Community	High (N=30)	0.25 (0.43)	0.43 (0.73)	0.25 (0.43)	0.71 (0.88)	0.40 (0.66)
		Medium (N=35)	1.00 (1.32)	0.89 (1.29)	0.67 (0.67)	1.56 (1.57)	1.03 (1.30)
		Low (N=32)	1.38 (0.86)	1.75 (0.97)	1.11 (1.10)	1.43 (1.05)	1.41 (1.03)
		Across Reading Level (N=97)	0.88 (1.05)	1.04 (1.17)	0.69 (0.87)	1.26 (1.29)	
Test 3, Part B	Knowledge of Definition of Community	High (N=30)	0.88 (0.33)	0.71 (0.45)	0.50 (0.50)	0.43 (0.49)	0.63 (0.48)
		Medium (N=35)	0.38 (0.48)	0.67 (0.47)	0.22 (0.42)	0.44 (0.50)	0.43 (0.49)
		Low (N=32)	0.50 (0.50)	0.00 (0.00)	0.44 (0.50)	0.29 (0.45)	0.31 (0.46)
		Across Reading Level (N=97)	0.58 (0.49)	0.46 (0.50)	0.38 (0.49)	0.39 (0.49)	

next page . . .

Test 4, Parts A and B	Knowledge of Interrelationships	High (N=30)	6.63 (0.86)	6.86 (1.36)	6.50 (1.80)	6.00 (0.93)	6.50 (1.34)
		Medium (N=35)	6.00 (0.87)	6.44 (1.42)	7.00 (1.41)	6.44 (2.54)	6.49 (1.73)
		Low (N=32)	5.25 (1.64)	4.88 (1.05)	4.33 (1.41)	5.00 (1.41)	4.84 (1.44)
		Across Reading Level (N=97)	5.96 (1.31)	6.04 (1.54)	5.92 (1.94)	5.87 (1.94)	
Test 4, Part C	Knowledge of Definitions of Population, Habitat, and Community	High (N=30)	1.50 (0.50)	0.86 (0.64)	0.88 (0.93)	1.00 (0.93)	1.07 (0.81)
		Medium (N=35)	0.50 (0.71)	0.67 (0.82)	0.67 (0.67)	0.22 (0.42)	0.51 (0.69)
		Low (N=32)	0.63 (0.99)	0.38 (0.48)	0.67 (0.82)	0.43 (0.73)	0.53 (0.79)
		Across Reading Level (N=97)	0.88 (0.88)	0.63 (0.70)	0.73 (0.81)	0.52 (0.77)	

[end of table]

Raw Score Means and Standard Deviations* for Main Study I (Retention)

Test	Dependent Variable	Previous Reading Achievement Level	Treatment Condition 1 (N=24)	2 (N=24)	3 (N=26)	4 (Control) (N=23)	Totals Across Condition
Test 1, Part A	Recognition of New Instances and Non-Instances of the Concept Population	High (N=30)	19.00 (3.16)	17.00 (2.51)	17.00 (2.65)	17.14 (2.75)	17.57 (2.92)
		Medium (N=35)	16.13 (2.85)	16.78 (1.62)	15.22 (2.86)	15.33 (1.70)	15.86 (2.40)
		Low (N=32)	15.88 (1.45)	15.13 (1.96)	15.56 (3.77)	12.86 (2.59)	14.94 (2.87)
		Across Reading Level (N=97)	17.00 (2.96)	16.29 (2.19)	15.88 (3.24)	15.13 (2.88)	
	Overgeneralization for the Concept Population	High (N=30)	4.38 (3.12)	7.29 (3.01)	6.38 (3.24)	5.00 (1.69)	5.73 (3.08)
		Medium (N=35)	6.38 (3.43)	6.00 (1.76)	8.11 (2.02)	6.56 (2.31)	6.77 (2.56)
		Low (N=32)	6.25 (1.92)	8.50 (2.87)	6.44 (2.71)	6.00 (2.39)	6.81 (2.70)
		Across Reading Level (N=97)	5.67 (3.04)	7.21 (2.77)	7.00 (2.80)	5.91 (2.26)	
	Undergeneralization for the Concept Population	High (N=30)	1.63 (1.93)	0.71 (1.39)	1.63 (2.12)	2.86 (2.47)	1.70 (2.15)
		Medium (N=35)	2.38 (2.23)	2.22 (2.20)	1.67 (2.45)	3.11 (2.64)	2.34 (2.45)
		Low (N=32)	2.88 (2.03)	1.38 (2.06)	2.89 (2.88)	5.71 (3.73)	3.13 (3.12)

next page . . .

*Note: Standard deviations are given in parentheses.

		Across Reading Level (N=97)	2.29 (2.13)	1.50 (2.04)	2.08 (2.59)	3.83 (3.23)	
Test 2, Part A	Recognition of New Instances and Non-Instances of the Concept Habitat	High (N=30)	18.63 (4.21)	17.14 (4.97)	16.75 (4.58)	18.57 (4.27)	17.77 (4.59)
		Medium (N=35)	16.13 (5.71)	12.44 (1.89)	14.22 (2.82)	14.67 (6.57)	14.31 (4.82)
		Low (N=32)	16.13 (3.82)	12.13 (2.85)	15.11 (6.14)	13.86 (3.94)	14.34 (4.68)
		Across Reading Level (N=97)	16.96 (4.80)	13.71 (4.02)	15.31 (4.83)	15.61 (5.58)	
	Overgeneralization for the Concept Habitat	High (N=30)	5.88 (4.62)	6.86 (5.11)	7.75 (4.79)	6.00 (4.50)	6.63 (4.82)
		Medium (N=35)	6.25 (4.84)	11.11 (2.18)	9.78 (3.49)	6.78 (4.29)	8.54 (4.31)
		Low (N=32)	5.75 (2.82)	9.50 (3.71)	7.67 (3.89)	10.14 (3.91)	8.19 (3.98)
		Across Reading Level (N=97)	5.96 (4.20)	9.33 (4.12)	8.42 (4.18)	7.57 (4.59)	
	Undergeneralization for the Concept Habitat	High (N=30)	0.50 (1.00)	1.00 (1.07)	0.50 (0.50)	0.29 (0.45)	0.57 (0.84)
		Medium (N=35)	2.63 (2.74)	1.44 (1.95)	1.00 (1.56)	3.56 (3.47)	2.14 (2.73)
		Low (N=32)	3.00 (2.65)	3.13 (3.26)	1.78 (2.57)	1.00 (0.93)	2.25 (2.68)
		Across Reading Level (N=97)	2.04 (2.52)	1.88 (2.47)	1.12 (1.87)	1.78 (2.67)	

Next page . . .

Test 3, Part A	Recognition of New Instances and Non-Instances of the Concept Community	High (N=30)	8.88 (1.27)	6.57 (2.38)	7.38 (1.93)	6.71 (1.48)	7.43 (2.03)
		Medium (N=35)	6.63 (3.12)	5.33 (1.25)	6.78 (1.31)	6.56 (2.31)	6.31 (2.19)
		Low (N=32)	5.38 (2.12)	5.13 (1.45)	6.22 (1.03)	5.00 (1.69)	5.47 (1.68)
		Across Reading Level (N=97)	6.96 (2.72)	5.63 (1.82)	6.77 (1.53)	6.13 (2.05)	
	Overgeneralization for the Concept Community	High (N=30)	1.00 (1.32)	2.71 (1.83)	2.00 (1.58)	2.29 (0.88)	1.97 (1.58)
		Medium (N=35)	2.63 (2.06)	3.56 (0.96)	2.78 (1.31)	2.22 (1.13)	2.80 (1.49)
		Low (N=32)	3.13 (1.36)	3.75 (1.48)	2.78 (0.63)	3.86 (0.99)	3.34 (1.24)
		Across Reading Level (N=97)	2.25 (1.85)	3.38 (1.49)	2.54 (1.28)	2.74 (1.26)	
	Undergeneralization for the Concept Community	High (N=30)	0.13 (0.33)	0.71 (1.03)	0.63 (0.86)	1.00 (0.93)	0.60 (0.88)
		Medium (N=35)	0.75 (1.64)	1.11 (1.10)	0.44 (0.56)	1.22 (1.40)	0.89 (1.26)
		Low (N=32)	1.50 (1.41)	1.00 (1.12)	1.00 (0.82)	1.14 (1.36)	1.16 (1.20)
		Across Reading Level (N=97)	0.79 (1.38)	0.96 (1.10)	0.69 (0.77)	1.13 (1.26)	
Tests 1,2, and 3 Parts B summed	Knowledge of Definitions of Population, Habitat, and Community	High (N=30)	1.50 (0.87)	1.29 (0.70)	1.50 (0.87)	1.86 (0.83)	1.53 (0.85)
		Medium (N=35)	1.00 (1.22)	1.22 (0.79)	0.89 (0.57)	1.33 (0.82)	1.11 (0.89)

next page . . .

		Low (N=32)	1.25 (0.97)	1.00 (0.50)	0.89 (0.57)	1.14 (0.64)	1.06 (0.70)
		Across Reading Level (N=97)	1.25 (1.05)	1.17 (0.69)	1.08 (0.73)	1.43 (0.82)	
Test 4, Parts A and B	Knowledge of Interrelationships	High (N=30)	7.13 (1.36)	7.29 (0.70)	7.25 (1.09)	6.43 (0.73)	7.03 (1.08)
		Medium (N=35)	6.25 (1.71)	6.67 (1.63)	6.44 (1.07)	6.67 (1.05)	6.51 (1.40)
		Low (N=32)	6.13 (1.45)	5.88 (1.54)	5.11 (1.29)	5.00 (1.77)	5.53 (1.58)
		Across Reading Level (N=97)	6.50 (1.58)	6.58 (1.50)	6.23 (1.45)	6.09 (1.44)	
Test 4, Part C	Knowledge of Definitions of Population, Habitat, and Community	High (N=30)	1.63 (0.99)	1.14 (0.99)	0.63 (0.70)	1.43 (1.29)	1.20 (1.08)
		Medium (N=35)	0.50 (0.50)	1.11 (0.57)	0.89 (0.87)	0.78 (0.63)	0.83 (0.70)
		Low (N=32)	0.50 (0.71)	0.50 (0.71)	0.67 (0.67)	0.43 (0.49)	0.53 (0.66)
		Across Reading Level (N=97)	0.88 (0.93)	0.92 (0.81)	0.73 (0.76)	0.87 (0.95)	

[end of table]

Raw Score Means and Standard Deviations* for Main Study II (Immediate Acquisitions)

Test	Dependent Variable	Previous Reading Achievement Level	Treatment Condition 1 (N=30)	2 (N=28)	3 (N=30)	4 (Control) (N=26)	Totals Across Condition
Test 1, Part A	Recognition of New Instances and Non-Instances of the Concept Population	High (N=40)	18.80 (2.36)	17.50 (2.58)	20.60 (2.46)	15.20 (2.44)	18.03 (3.15)
		Medium (N=38)	18.10 (1.87)	17.10 (3.81)	17.30 (1.90)	12.50 (2.96)	16.45 (3.44)
		Low (N=36)	13.10 (2.66)	15.63 (3.24)	14.70 (3.49)	13.00 (2.06)	14.08 (3.13)
		Across Reading Level (N=114)	16.67 (3.44)	16.82 (3.34)	17.53 (3.62)	13.69 (2.78)	
	Overgeneralization for the Concept Population	High (N=40)	3.10 (1.97)	5.20 (2.32)	4.20 (2.71)	6.10 (1.51)	4.65 (2.44)
		Medium (N=38)	3.60 (2.06)	4.70 (2.05)	6.50 (1.96)	5.50 (1.80)	5.05 (2.26)
		Low (N=36)	5.60 (1.91)	5.25 (2.11)	5.40 (2.15)	6.00 (2.29)	5.56 (2.13)
		Across Reading Level (N=114)	4.10 (2.26)	5.04 (2.18)	5.37 (2.48)	5.88 (1.89)	
	Undergeneralization for the Concept Population	High (N=40)	3.10 (1.92)	2.20 (2.04)	0.20 (0.60)	3.70 (2.65)	2.30 (2.36)
		Medium (N=38)	3.10 (1.58)	3.20 (2.68)	1.20 (2.75)	7.00 (3.57)	3.45 (3.35)

next page . . .

*Note: Standard deviations are given in parentheses.

		Low (N=36)	6.20 (3.79)	4.13 (3.22)	4.90 (4.09)	6.00 (2.29)	5.33 (3.57)
		Across Reading Level (N=114)	4.13 (3.00)	3.11 (2.76)	2.10 (3.51)	5.42 (3.20)	
Test 1 Part B	Knowledge of Definition of Population	High (N=40)	0.70 (0.46)	0.90 (0.30)	0.80 (0.40)	0.10 (0.30)	0.63 (0.48)
		Medium (N=38)	0.20 (0.40)	0.80 (0.40)	0.80 (0.40)	0.13 (0.33)	0.50 (0.50)
		Low (N=36)	0.10 (0.30)	0.25 (0.43)	0.30 (0.46)	0.00 (0.00)	0.17 (0.37)
		Across Reading Level (N=114)	0.33 (0.47)	0.68 (0.47)	0.63 (0.48)	0.08 (0.27)	
Test 2, Part A	Recognition of New Instances and Non-Instances of the Concept Habitat	High (N=40)	20.50 (2.50)	20.70 (3.10)	22.20 (2.48)	17.70 (3.61)	20.28 (3.38)
		Medium (N=38)	18.40 (2.84)	15.90 (7.48)	17.50 (4.57)	15.63 (4.44)	16.92 (5.27)
		Low (N=36)	11.90 (3.94)	17.75 (6.16)	14.60 (4.52)	11.75 (1.20)	13.92 (4.93)
		Across Reading Level (N=114)	16.93 (4.83)	18.14 (6.20)	18.10 (5.06)	15.23 (4.20)	
	Overgeneralization for the Concept Habitat	High (N=40)	3.00 (2.28)	3.30 (2.45)	2.30 (2.15)	5.80 (4.49)	3.60 (3.28)
		Medium (N=38)	4.40 (2.25)	4.20 (4.64)	5.20 (2.82)	8.13 (4.75)	5.34 (4.01)
		Low (N=36)	9.50 (3.14)	5.50 (5.27)	8.60 (4.29)	11.75 (1.92)	8.86 (4.40)

next page . . .

		Across Reading Level (N=114)	5.63 (3.81)	4.25 (4.31)	5.37 (4.12)	8.35 (4.68)	
	Undergeneralization for the Concept Habitat	High (N=40)	1.50 (1.43)	1.00 (1.41)	0.50 (0.67)	1.50 (1.69)	1.13 (1.42)
		Medium (N=38)	2.20 (2.04)	4.90 (3.36)	2.20 (2.56)	1.25 (1.92)	2.71 (2.90)
		Low (N=36)	3.60 (2.29)	1.75 (2.38)	1.80 (1.54)	1.50 (1.50)	2.22 (2.15)
		Across Reading Level (N=114)	2.43 (2.14)	2.61 (3.06)	1.50 (1.91)	1.42 (1.71)	
Test 2, Part B	Knowledge of Definition of Habitat	High (N=40)	0.50 (0.50)	0.90 (0.30)	0.90 (0.30)	0.60 (0.49)	0.73 (0.45)
		Medium (N=38)	0.80 (0.40)	0.70 (0.46)	0.90 (0.30)	0.50 (0.50)	0.74 (0.44)
		Low (N=36)	0.40 (0.49)	0.88 (0.33)	0.50 (0.50)	0.63 (0.48)	0.58 (0.49)
		Across Reading Level (N=114)	0.57 (0.50)	0.82 (0.38)	0.77 (0.42)	0.58 (0.49)	
Test 3, Part A	Recognition of New Instances and Non-Instances of the Concept Community	High (N=40)	9.00 (1.18)	8.80 (1.08)	8.80 (1.40)	7.60 (1.56)	8.55 (1.43)
		Medium (N=38)	7.10 (2.47)	7.40 (1.43)	7.70 (1.68)	6.13 (1.69)	7.13 (1.95)
		Low (N=36)	5.10 (1.14)	7.50 (2.18)	5.90 (2.17)	5.13 (2.26)	5.86 (2.18)
		Across Reading Level (N=114)	7.07 (2.34)	7.93 (1.71)	7.47 (2.14)	6.38 (2.11)	

next page . . .

	Overgeneralization for the Concept Community	High (N=40)	0.60 (1.20)	0.70 (0.90)	0.80 (1.08)	1.60 (1.50)	0.93 (1.25)
		Medium (N=38)	1.70 (1.27)	1.10 (1.30)	1.70 (1.27)	3.00 (1.41)	1.82 (1.47)
		Low (N=36)	3.50 (0.92)	1.88 (1.83)	2.70 (1.27)	3.38 (1.32)	2.89 (1.49)
		Across Reading Level (N=114)	1.93 (1.65)	1.18 (1.44)	1.73 (1.44)	2.58 (1.62)	
	Undergeneralization for the Concept Community	High (N=40)	0.40 (0.66)	0.50 (0.67)	0.40 (0.66)	0.80 (0.87)	0.53 (0.74)
		Medium (N=38)	1.20 (1.33)	1.50 (1.12)	0.60 (0.80)	0.88 (1.36)	1.05 (1.21)
		Low (N=36)	1.40 (0.80)	0.63 (0.86)	1.40 (1.28)	1.50 (1.58)	1.25 (1.21)
		Across Reading Level (N=114)	1.00 (1.06)	0.89 (1.01)	0.80 (1.05)	1.04 (1.32)	
Test 3, Part B	Knowledge of Definition of Community	High (N=40)	0.50 (0.50)	0.90 (0.30)	0.70 (0.46)	0.60 (0.49)	0.68 (0.47)
		Medium (N=38)	0.50 (0.50)	0.90 (0.30)	0.80 (0.40)	0.38 (0.48)	0.66 (0.47)
		Low (N=36)	0.50 (0.50)	0.63 (0.48)	0.40 (0.49)	0.25 (0.43)	0.44 (0.50)
		Across Reading Level (N=114)	0.50 (0.50)	0.82 (0.38)	0.63 (0.48)	0.42 (0.49)	
		next page . . .					

Test 4, Parts A and B	Knowledge of Interrelationships	High (N=40)	6.70 (0.78)	6.60 (1.02)	6.50 (1.28)	7.00 (1.00)	6.70 (1.05)
		Medium (N=38)	6.60 (1.62)	6.50 (1.28)	6.30 (1.27)	7.13 (0.78)	6.61 (1.33)
		Low (N=36)	5.60 (1.36)	7.00 (0.71)	5.60 (1.50)	5.88 (1.36)	5.97 (1.40)
		Across Reading Level (N=114)	6.30 (1.39)	6.68 (1.07)	6.13 (1.41)	6.69 (1.20)	
Test 4 Part C	Knowledge of Definitions of Population, Habitat, and Community	High (N=40)	1.60 (0.80)	2.00 (0.77)	2.00 (1.18)	1.40 (1.11)	1.75 (1.02)
		Medium (N=38)	1.50 (1.12	1.2C (1.17)	1.60 (1.02)	1.50 (0.87)	1.45 (1.07)
		Low (N=36)	0.60 (0.66)	0.88 (0.78)	0.80 (0.87)	0.38 (0.70)	0.67 (0.78)
		Across Reading Level (N=114)	1.23 (0.99)	1.39 (1.05)	1.47 (1.15)	1.12 (1.05)	

[end of table]

Raw Score Means and Standard Deviations* for Main Study II (Retention)

Test	Dependent Variable	Previous Reading Achievement Level	Treatment Condition				Totals Across Condition
			1 (N=30)	2 (N=28)	3 (N=30)	4 (Control) (N=26)	
Test 1, Part A	Recognition of New Instances and Non-Instances of the Concept Population	High (N=40)	17.10 (1.97)	18.70 (3.29)	21.40 (3.32)	15.80 (2.23)	18.25 (3.47)
		Medium (N=38)	16.60 (2.50)	16.70 (4.12)	17.60 (4.10)	11.88 (2.52)	15.89 (4.04)
		Low (N=36)	12.30 (3.23)	17.00 (3.04)	14.10 (4.81)	12.13 (1.83)	13.81 (3.96)
		Across Reading Level (N=114)	15.33 (3.39)	17.50 (3.66)	17.70 (5.09)	13.46 (2.89)	
	Overgeneralization for the Concept Population	High (N=40)	4.70 (2.28)	4.90 (2.88)	3.60 (3.32)	6.30 (1.95)	4.88 (2.83)
		Medium (N=38)	4.60 (2.01)	4.80 (1.94)	5.50 (2.54)	5.75 (1.98)	5.13 (2.19)
		Low (N=36)	6.80 (1.40)	4.38 (2.34)	5.80 (2.44)	7.13 (3.41)	6.06 (2.66)
		Across Reading Level (N=114)	5.37 (2.18)	4.71 (2.43)	4.97 (2.96)	6.38 (2.56)	
	Undergeneralization for the Concept Population	High (N=40)	3.20 (1.99)	1.30 (1.35)	0.00 (0.00)	2.90 (2.47)	1.85 (2.15)
		Medium (N=38)	3.80 (1.72)	3.50 (3.23)	1.80 (3.43)	7.38 (4.30)	3.95 (3.78)
		Low (N=36)	5.90 (3.70)	3.63 (4.03)	5.10 (4.70)	5.75 (3.60)	5.14 (4.14)

next page . . .

*Note: Standard deviations are given in parentheses.

		Across Reading Level (N=114)	4.30 (2.87)	2.75 (3.19)	2.30 (3.97)	5.15 (3.95)	
Test 2, Part A	Recognition of New Instances and Non-Instances of the Concept Habitat	High (N=40)	20.20 (3.49)	20.80 (3.60)	22.40 (1.96)	18.30 (4.05)	20.43 (3.67)
		Medium (N=38)	15.90 (5.03)	15.20 (6.42)	16.90 (3.36)	15.50 (4.44)	15.89 (5.00)
		Low (N=36)	12.20 (2.18)	17.25 (4.35)	13.20 (2.96)	12.38 (0.99)	13.64 (3.47)
		Across Reading Level (N=114)	16.10 (4.98)	17.79 (5.52)	17.50 (4.72)	15.62 (4.32)	
	Overgeneralization for the Concept Habitat	High (N=40)	3.50 (3.72)	3.70 (3.69)	2.40 (1.96)	5.50 (4.10)	3.78 (3.64)
		Medium (N=38)	6.60 (4.27)	6.00 (5.04)	6.00 (2.49)	7.13 (5.16)	6.39 (4.35)
		Low (N=36)	8.70 (3.07)	7.25 (4.74)	10.30 (3.03)	12.00 (1.66)	9.56 (3.69)
		Across Reading Level (N=114)	6.27 (4.29)	5.54 (4.74)	6.23 (4.10)	8.00 (4.80)	
	Undergeneralization for the Concept Habitat	High (N=40)	1.20 (1.17)	0.40 (0.66)	0.20 (0.40)	1.20 (1.54)	0.75 (1.13)
		Medium (N=38)	2.50 (2.29)	3.70 (3.13)	2.10 (2.66)	2.38 (2.96)	2.68 (2.84)
		Low (N=36)	4.10 (2.95)	0.50 (0.71)	1.50 (1.63)	0.63 (0.99)	1.81 (2.38)
		Across Reading Level (N=114)	2.60 (2.55)	1.61 (2.50)	1.27 (1.98)	1.38 (2.10)	

next page . . .

Test 3, Part A	Recognition of New Instances and Non-Instances of the Concept Community	High (N=40)	8.60 (1.02)	9.00 (1.18)	8.60 (1.69)	7.90 (1.58)	8.53 (1.45)
		Medium(N=38)	6.80 (1.40)	7.50 (1.20)	6.70 (1.85)	6.50 (1.00)	6.89 (1.47)
		Low (N=36)	5.20 (1.25)	7.88 (1.69)	5.70 (1.49)	5.50 (1.32)	6.00 (1.76)
		Across Reading Level (N=114)	6.87 (1.86)	8.14 (1.51)	7.00 (2.07)	6.73 (1.68)	
	Overgeneralization for the Concept Community	High (N=40)	0.90 (0.94)	0.80 (1.25)	0.80 (1.25)	2.10 (1.58)	1.15 (1.39)
		Medium (N=38)	1.90 (0.94)	1.50 (1.20)	2.00 (1.26)	2.63 (1.41)	1.97 (1.27)
		Low (N=36)	3.30 (1.27)	1.63 (1.65)	3.40 (1.28)	4.00 (1.00)	3.11 (1.56)
		Across Reading Level (N=114)	2.03 (1.45)	1.29 (1.41)	2.07 (1.65)	2.85 (1.59)	
	Undergeneralization for the Concept Community	High (N=40)	0.40 (0.80)	0.20 (0.40)	0.60 (0.80)	0.00 (0.00)	0.30 (0.64)
		Medium (N=38)	1.30 (1.01)	1.00 (0.89)	1.30 (1.19)	0.88 (0.93)	1.13 (1.03)
		Low (N=36)	1.50 (0.67)	0.50 (0.71)	0.90 (0.54)	0.50 (0.71)	0.89 (0.77)
		Across Reading Level (N=114)	1.07 (0.96)	0.57 (0.78)	0.93 (0.93)	0.42 (0.74)	
		next page . . .					

Tests 1,2, and 3 Parts B summed	Knowledge of Definitions of Population, Habitat, and Community	High (N=40)	2.00 (1.00)	2.10 (1.22)	2.60 (0.66)	1.50 (0.81)	2.05 (1.02)
		Medium (N=38)	1.50 (0.67)	2.10 (0.83)	2.40 (0.66)	1.25 (0.83)	1.84 (0.87)
		Low (N=36)	1.10 (0.83)	2.13 (0.93)	1.40 (1.20)	1.13 (0.93)	1.42 (1.06)
		Across Reading Level (N=114)	1.53 (0.92)	2.11 (1.01)	2.13 (1.02)	1.31 (0.87)	
Test 4, Parts A and B	Knowledge of Interrelation-ships	High (N=40)	5.40 (2.33)	6.50 (0.81)	6.60 (1.20)	7.10 (0.94)	6.40 (1.58)
		Medium (N=38)	6.70 (1.19)	5.50 (2.16)	6.60 (1.28)	5.63 (1.58)	6.13 (1.69)
		Low (N=36)	5.50 (1.50)	6.75 (1.20)	5.00 (1.73)	6.00 (1.73)	5.75 (1.69)
		Across Reading Level (N=114)	5.87 (1.84)	6.21 (1.61)	6.07 (1.61)	6.31 (1.56)	
Test 4, Part C	Knowledge of Definitions of Population, Habitat, and Community	High (N=40)	1.20 (1.25)	1.80 (0.75)	2.00 (1.26)	1.40 (1.11)	1.60 (1.16)
		Medium (N=38)	1.00 (1.00)	1.00 (0.63)	1.50 (1.02)	1.25 (0.83)	1.18 (0.91)
		Low (N=36)	0.60 (0.66)	1.00 (1.12)	0.70 (0.90)	0.50 (0.71)	0.69 (0.88)
		Across Reading Level (N=114)	0.93 (1.03)	1.29 (0.92)	1.40 (1.20)	1.08 (1.00)	

[end of table]

Raw Score Means and Standard Deviations* for Main Study III (Immediate Acquisitions)

Test	Dependent Variable	Previous Reading Achievement Level	Treatment Condition				Totals Across Condition
			1 (N=24)	2 (N=26)	3 (N=26)	4 (Control) (N=26)	
Test 1, Part A	Recognition of New Instances and Non-Instances of the Concept Population	High (N=34)	21.38 (2.06)	18.63 (2.50)	20.56 (2.27)	13.89 (1.85)	18.53 (3.66)
		Medium (N=35)	17.00 (3.86)	17.00 (3.09)	17.33 (5.08)	12.63 (2.60)	16.09 (4.25)
		Low (N=33)	15.71 (2.43)	15.56 (1.89)	14.75 (2.77)	12.44 (2.79)	14.55 (2.83)
		Across Reading Level (N=102)	18.08 (3.80)	17.00 (2.83)	17.65 (4.31)	13.00 (2.53)	
	Overgeneralization for the Concept Population	High (N=34)	2.88 (2.26)	6.38 (2.50)	4.44 (2.27)	6.78 (1.03)	5.15 (2.59)
		Medium (N=35)	5.22 (2.44)	7.33 (2.11)	5.67 (2.91)	4.63 (2.39)	5.74 (2.68)
		Low (N=33)	5.43 (1.29)	7.11 (2.33)	7.38 (1.22)	6.67 (1.83)	6.70 (1.90)
		Across Reading Level (N=102)	4.50 (2.40)	6.96 (2.34)	5.77 (2.56)	6.08 (2.06)	
	Undergeneralization for the Concept Population	High (N=34)	0.75 (0.97)	0.00 (0.00)	0.00 (0.00)	4.33 (1.76)	1.32 (2.10)
		Medium (N=35)	2.67 (2.36)	0.67 (1.56)	2.00 (3.09)	7.63 (3.90)	3.11 (3.82)
		next page . . .					

*Note: Standard deviations are given in parentheses.

		Low (N=33)	3.86 (1.88)	2.33 (1.94)	2.50 (2.35)	5.56 (1.57)	3.58 (2.36)
		Across Reading Level (N=102)	2.38 (2.23)	1.04 (1.76)	1.46 (2.48)	5.77 (2.90)	
Test 1, Part B	Knowledge of Definition of Population	High (N=34)	0.88 (0.33)	0.75 (0.43)	0.67 (0.47)	0.56 (0.50)	0.71 (0.46)
		Medium (N=35)	0.44 (0.50)	0.67 (0.47)	0.78 (0.42)	0.38 (0.48)	0.57 (0.49)
		Low (N=33)	0.29 (0.45)	0.33 (0.47)	0.63 (0.48)	0.56 (0.50)	0.45 (0.50)
		Across Reading Level (N=102)	0.54 (0.50)	0.58 (0.49)	0.69 (0.46)	0.50 (0.50)	
Test 2, Part A	Recognition of New Instances and Non-Instances of the Concept Habitat	High (N=34)	23.13 (2.20)	14.50 (4.15)	18.56 (3.65)	15.56 (3.69)	17.88 (4.80)
		Medium (N=35)	18.00 (5.01)	14.11 (3.98)	15.11 (5.40)	11.63 (2.34)	14.80 (4.94)
		Low (N=33)	16.14 (4.76)	12.56 (3.20)	12.50 (2.45)	11.89 (1.45)	13.12 (3.49)
		Across Reading Level (N=102)	19.17 (5.10)	13.69 (3.88)	15.50 (4.76)	13.08 (3.22)	
	Overgeneralization for the Concept Habitat	High (N=34)	1.13 (1.96)	10.00 (4.21)	6.33 (3.46)	8.11 (3.96)	6.44 (4.77)
		Medium (N=35)	5.67 (4.59)	9.78 (3.55)	6.00 (4.45)	11.38 (1.49)	8.11 (4.48)
		Low (N=33)	6.71 (4.43)	11.44 (2.71)	9.00 (2.50)	10.22 (2.44)	9.52 (3.49)

next page . . .

		Across Reading Level (N=102)	4.46 (4.55)	10.42 (3.60)	7.04 (3.83)	9.85 (3.16)	
	Undergeneralization for the Concept Habitat	High (N=34)	0.75 (0.83)	0.50 (0.71)	0.11 (0.31)	1.11 (1.37)	0.62 (0.97)
		Medium (N=35)	1.33 (1.33)	1.11 (1.20)	1.78 (1.81)	2.00 (2.96)	1.54 (1.95)
		Low (N=33)	2.00 (2.62)	1.00 (1.15)	3.50 (2.18)	2.78 (2.25)	2.30 (2.29)
		Across Reading Level (N=102)	1.33 (1.77)	0.88 (1.09)	1.73 (2.12)	1.96 (2.36)	
Test 2, Part B	Knowledge of Definition of Habitat	High (N=34)	1.00 (0.00)	0.88 (0.33)	1.00 (0.00)	0.67 (0.47)	0.88 (0.32)
		Medium (N=35)	0.67 (0.47)	0.78 (0.42)	0.78 (0.42)	0.25 (0.43)	0.63 (0.48)
		Low (N=33)	0.29 (0.45)	0.44 (0.50)	0.50 (0.50)	0.33 (0.47)	0.39 (0.49)
		Across Reading Level (N=102)	0.67 (0.47)	0.69 (0.46)	0.77 (0.42)	0.42 (0.49)	
Test 3, Part A	Recognition of New Instances and Non-Instances of the Concept Community	High (N=34)	9.00 (1.66)	6.38 (1.73)	8.78 (1.87)	7.00 (1.56)	7.79 (2.04)
		Medium (N=35)	6.89 (2.02)	7.56 (1.71)	7.56 (2.06)	5.75 (1.09)	6.97 (1.92)
		Low (N=33)	6.57 (0.90)	6.00 (1.25)	6.63 (1.93)	5.56 (1.77)	6.15 (1.60)
		Across Reading Level (N=102)	7.50 (1.96)	6.65 (1.71)	7.69 (2.14)	6.12 (1.65)	

next page . . .

	Overgeneralization for the Concept Community	High (N=34)	0.75 (1.09)	3.13 (1.69)	0.78 (1.13)	2.33 (1.33)	1.74 (1.67)
		Medium (N=35)	2.00 (1.25)	2.22 (1.62)	2.22 (1.75)	2.88 (0.78)	2.31 (1.45)
		Low (N=33)	1.86 (0.99)	3.33 (0.94)	2.25 (1.48)	2.89 (1.29)	2.64 (1.32)
		Across Reading Level (N=102)	1.54 (1.26)	2.88 (1.53)	1.73 (1.63)	2.69 (1.20)	
	Undergeneralization for the Concept Community	High (N=34)	0.25 (0.66)	0.50 (0.71)	0.44 (0.96)	0.67 (0.94)	0.47 (0.85)
		Medium (N=35)	1.11 (1.37)	0.22 (0.42)	0.22 (0.63)	1.38 (0.99)	0.71 (1.06)
		Low (N=33)	1.57 (1.05)	0.67 (0.82)	1.13 (0.93)	1.56 (1.07)	1.21 (1.04)
		Across Reading Level (N=102)	0.96 (1.21)	0.46 (0.69)	0.58 (0.93)	1.19 (1.07)	
Test 3, Part B	Knowledge of Definition of Community	High (N=34)	0.88 (0.33)	0.75 (0.43)	0.89 (0.31)	0.44 (0.50)	0.74 (0.44)
		Medium (N=35)	0.56 (0.50)	0.67 (0.47)	0.78 (0.42)	0.13 (0.33)	0.54 (0.50)
		Low (N=33)	0.71 (0.45)	0.44 (0.50)	0.63 (0.48)	0.11 (0.31)	0.45 (0.50)
		Across Reading Level (N=102)	0.71 (0.45)	0.62 (0.49)	0.77 (0.42)	0.23 (0.42)	
		next page . . .					

Test 4, Parts A and B	Knowledge of Interrelationships	High (N=34)	6.75 (1.09)	5.75 (1.39)	6.78 (1.23)	6.56 (1.26)	6.47 (1.31)
		Medium (N=35)	4.78 (2.20)	7.00 (1.15)	6.33 (1.76)	5.25 (1.48)	5.86 (1.91)
		Low (N=33)	5.29 (1.98)	4.56 (1.89)	4.50 (1.50)	5.11 (2.18)	4.85 (1.94)
		Across Reading Level (N=102)	5.58 (2.02)	5.77 (1.83)	5.92 (1.80)	5.65 (1.82)	
Test 4, Part C	Knowledge of Definitions of Population, Habitat, and Community	High (N=34)	2.63 (0.70)	1.13 (1.05)	2.33 (0.67)	1.11 (1.10)	1.79 (1.13)
		Medium (N=35)	0.78 (0.63)	1.11 (0.57)	1.11 (0.99)	0.13 (0.33)	0.80 (0.79)
		Low (N=33)	0.86 (0.99)	1.44 (0.96)	0.88 (1.05)	0.67 (0.82)	0.97 (1.00)
		Across Reading Level (N=102)	1.42 (1.15)	1.23 (0.89)	1.46 (1.12)	0.65 (0.92)	

[end of table]

Raw Score Means and Standard Deviations* for Main Study III (Retention)

Test	Dependent Variable	Previous Reading Achievement Level	Treatment Condition				Totals Across Condition
			1 (N=24)	2 (N=26)	3 (N=26)	4 (Control) (N=26)	
Test 1, Part A	Recognition of New Instances and Non-Instances of the Concept Population	High (N=34)	22.00 (2.55)	18.38 (1.87)	20.00 (3.43)	16.00 (1.76)	19.03 (3.34)
		Medium (N=35)	16.11 (5.02)	17.11 (1.66)	17.44 (3.89)	14.00 (2.55)	16.23 (3.78)
		Low (N=33)	15.14 (3.52)	14.67 (3.30)	15.75 (2.17)	10.56 (3.89)	13.91 (3.90)
		Across Reading Level (N=102)	17.79 (4.92)	16.65 (2.85)	17.81 (3.71)	13.50 (3.68)	
	Overgeneralization for the Concept Population	High (N=34)	2.88 (2.47)	6.63 (1.87)	4.89 (3.31)	5.67 (2.67)	5.03 (2.98)
		Medium (N=35)	5.11 (3.11)	7.33 (1.25)	5.11 (2.18)	5.75 (3.73)	5.83 (2.85)
		Low (N=33)	7.00 (1.41)	7.67 (2.05)	6.75 (1.30)	8.44 (2.36)	7.52 (1.99)
		Across Reading Level (N=102)	4.92 (2.98)	7.23 (1.80)	5.54 (2.58)	6.65 (3.22)	
	Undergeneralization for the Concept Population	High (N=34)	0.13 (0.33)	0.00 (0.00)	0.00 (0.00)	3.33 (2.40)	0.91 (1.92)
		Medium (N=35)	3.67 (3.37)	0.56 (1.26)	2.44 (3.37)	5.25 (3.63)	2.91 (3.48)
		Low (N=33)	2.86 (3.04)	2.67 (2.75)	2.38 (2.69)	5.00 (3.27)	3.27 3.14

next page . . .

*Note: Standard deviations are given in parentheses.

		Across Reading Level (N=102)	2.25 (3.06)	1.12 (2.12)	1.58 (2.73)	4.50 (3.24)	
Test 2, Part A	Recognition of New Instances and Non-Instances of the Concept Habitat	High (N=34)	21.25 (5.61)	14.50 (3.50)	19.33 (4.14)	15.78 (4.85)	17.71 (5.30)
		Medium (N=35)	16.33 (4.83)	12.89 (3.98)	16.78 (5.43)	13.88 (3.22)	15.00 (4.77)
		Low (N=33)	15.57 (4.53)	11.89 (2.85)	11.88 (4.37)	11.78 (0.92)	12.64 (3.71)
		Across Reading Level (N=102)	17.75 (5.61)	13.04 (3.63)	16.15 (5.59)	13.81 (3.79)	
	Overgeneralization for the Concept Habitat	High (N=34)	2.50 (3.24)	10.00 (3.54)	5.67 (4.14)	7.78 (4.31)	6.50 (4.70)
		Medium (N=35)	6.33 (3.83)	11.00 (3.40)	5.78 (3.58)	9.50 (3.74)	8.11 (4.25)
		Low (N=33)	7.00 (2.88)	10.22 (3.19)	9.25 (2.44)	10.44 (2.75)	9.36 (3.12)
		Across Reading Level (N=102)	5.25 (3.91)	10.42 (3.40)	6.81 (3.85)	9.23 (3.83)	
	Undergeneralization for the Concept Habitat	High (N=34)	1.25 (2.59)	0.50 (0.87)	0.00 (0.00)	1.44 (1.89)	0.79 (1.75)
		Medium (N=35)	2.33 (2.75)	1.11 (0.99)	2.33 (2.11)	1.63 (2.18)	1.86 (2.17)
		Low (N=33)	2.29 (2.91)	2.22 (2.20)	3.88 (3.14)	2.78 (3.42)	2.79 (3.02)
		Across Reading Level (N=102)	1.96 (2.79)	1.31 (1.66)	2.00 (2.66)	1.96 (2.67)	

next page . . .

Test 3, Part A	Recognition of New Instances and Non-Instances of the Concept Community	High (N=34)	8.50 (2.29)	7.50 (1.94)	8.56 (2.22)	7.11 (2.02)	7.91 (2.21)
		Medium (N=35)	5.56 (2.01)	6.33 (1.56)	8.00 (2.21)	5.63 (0.99)	6.40 (2.03)
		Low (N=33)	5.29 (2.12)	5.44 (1.64)	5.25 (0.83)	5.33 (1.41)	5.33 (1.55)
		Across Reading Level (N=102)	6.46 (2.58)	6.38 (1.90)	7.35 (2.37)	6.04 (1.74)	
	Overgeneralization for the Concept Community	High (N=34)	0.88 (1.36)	2.25 (1.85)	1.22 (1.69)	2.33 (1.63)	1.68 (1.76)
		Medium (N=35)	2.11 (1.37)	3.33 (1.94)	1.33 (1.33)	3.50 (1.41)	2.54 (1.78)
		Low (N=33)	2.36 (1.36)	3.56 (1.26)	3.13 (0.93)	3.22 (1.62)	3.21 (1.34)
		Across Reading Level (N=102)	1.92 (1.58)	3.08 (1.80)	1.85 (1.61)	3.00 (1.64)	
	Undergeneralization for the Concept Community	High (N=34)	0.63 (0.99)	0.25 (0.43)	0.22 (0.63)	0.56 (0.83)	0.41 (0.77)
		Medium (N=35)	2.33 (1.33)	0.33 (0.67)	0.67 (1.05)	0.88 (0.78)	1.06 (1.26)
		Low (N=33)	1.86 (1.46)	1.00 (1.05)	1.63 (0.99)	1.44 (1.26)	1.45 (1.23)
		Across reading Level (N=102)	1.63 (1.47)	0.54 (0.84)	0.81 (1.07)	0.96 (1.06)	

next page . . .

Tests 1, 2, 3 Parts B summed	Knowledge of Definitions of Population, Habitat, and Community	High (N=34)	2.63 (0.70)	2.13 (1.05)	2.78 (0.42)	1.78 (0.79)	2.32 (0.87)
		Medium (N=35)	1.44 (1.17)	1.78 (1.13)	2.00 (1.05)	1.38 (0.70)	1.66 (1.07)
		Low (N=33)	0.86 (0.99)	1.33 (0.94)	0.75 (0.83)	0.78 (0.63)	0.94 (0.89)
		Across Reading Level (N=102)	1.67 (1.21)	1.73 (1.09)	1.88 (1.15)	1.31 (0.82)	
Test 4, Parts A and B	Knowledge of Interrelation-ships	High (N=34)	6.88 (1.05)	6.50 (1.41)	7.44 (0.50)	7.22 (0.92)	7.03 (1.07)
		Medium (N=35)	5.67 (1.41)	6.00 (1.76)	6.33 (1.76)	6.38 (0.86)	6.09 (1.54)
		Low (N=33)	5.57 (1.40)	6.11 (1.37)	4.75 (1.56)	5.44 (1.50)	5.48 (1.5[illegible]
		Across Reading Level (N=102)	6.04 (1.43)	6.19 (1.54)	6.23 (1.76)	6.35 (1.36)	
Test 4, Part C	Knowledge of Definitions of Population, Habitat, and Community	High (N=34)	2.25 (0.83)	1.13 (0.78)	1.78 (1.03)	0.89 (0.74)	1.50 (1.01)
		Medium (N=35)	1.33 (1.33)	1.11 (0.99)	0.89 (1.20)	0.38 (0.48)	0.94 (1.12)
		Low (N=33)	1.29 (1.16)	0.78 (0.63)	0.38 (0.48)	0.89 (0.87)	0.82 (0.87)
		Across Reading Level (N=102)	1.63 (1.22)	1.00 (0.83)	1.04 (1.13)	0.73 (0.76)	

[end of table]