This is to certify that the

thesis entitled
AN EVALUATION OF INTERPRETIVE PROGRAMS AND RESEARCH METHODS AT INDIANA DUNES NATIONAL LAKESHORE: A PILOT STUDY

presented by

Charlotte Francis Young

has been accepted towards fulfillment
of the requirements for

Master of Science degree in Park and Recreation Resources

Maureen K. McDonough

Major professor
Maureen McDonough

Date 5/3/84

AN EVALUATION OF INTERPRETIVE PROGRAMS AND RESEARCH METHODS AT INDIANA DUNES NATIONAL LAKESHORE: A PILOT STUDY

By

Charlotte Francis Young

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Park and Recreation Resources

1984

ABSTRACT

AN EVALUATION OF INTERPRETIVE PROGRAMS
AND RESEARCH METHODS
AT INDIANA DUNES NATIONAL LAKESHORE:
A PILOT STUDY

By

Charlotte Francis Young

Effective interpretation attempts to modify visitors' cognitive, affective and psychomotor behaviors toward a resource. Several methods have been used to ascertain if interpretation changes these types of behaviors. This study's two objectives were to evaluate the effectiveness of Indiana Dunes National Lakeshore's dune succession interpretive programs and to investigate four methods (paper and pencil games, questionnaires, behavior intention stories and observations) used for evaluating interpretive program effectiveness.

Program effectiveness was operationally defined as: 1) extent to which prestipulated program objectives were met; 2) changes in participants' cognitive, affective and psychomotor behaviors; and 3) changes in participants' schematic differentiation about dunes. For this first study objective, 112 fourth to sixth graders served as the sample for a pretest-posttest-control group experiment. The methods were correlated to determine if they could be used interchangeably to measure "environmental behavior," a trait which was defined to contain cognitive, affective and psychomotor behavioral components. Twenty-seven

students participated in this second study objective.

Several trends were identified from the data. The prestipulated program objectives were not attained. Participants were knowledgeable about proper behavior such as protecting plants and animals, and had desired attitudes before the program. The greatest change in cognitive behavior was knowing what a dune was. Nearly all participants indicated they enjoyed the program. Generally, participants did not stay on the trail and they most often put trash in their pocket rather than a trash can. Correlation analysis revealed that all methods (paper and pencil games, questionnaires, behavior intention stories and observation) must be used to measure the "environmental behavior" trait.

ACKNOWLEDGMENTS

I would like to thank many people who helped me complete this research project. First, I express my appreciation to my committee members, Dr. Joe Fridgen, Mr. Glenn Dudderar and especially my major professor, Dr. Maureen McDonough, for their endless hours of discussing, critiquing and reading this project.

After a "long and bumpy road," I wish to thank Eastern Parks and Monuments Association for making this project possible. The students and teachers who participated in this study also deserve thanks for their cooperation and patience.

I greatly appreciate the long hours that Francie Freese spent typing this document.

Last, but certainly not least, I thank my husband, Peter Brandon, for his tolerance, understanding, support and helpful suggestions during every phase of this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Interpretation is a specialized type of communication that began as "nature guiding" in the early 1900's (Sharpe, 1982). Numerous authors have proposed more specific definitions. For instance, Tilden (1977) defined interpretation as an educational activity which aims to reveal meanings and relationships through the use of original objects, by first-hand experience, and by illustrative media, rather than simple communication of facts. Cherem (1975) suggested another definition: "(in part) the artful ability to make an environment or subject matter come to life for a particular group of visitors." Peart (1978) has devised the most comprehensive definition: "interpretation is any communication process designed to reveal meanings and relationships of our cultural and natural heritage to the public (primarily) through first hand involvement with an object, artifact, landscape of site." These definitions all include several essential aspects. Interpretation is fun for visitors and relates to their experiences. The communicator or interpreter attempts to "interpret" difficult, often scientific concepts into terms that the audience understands and encourages the audience to discover information themselves. Interpreters try to use areas where the audience can be directly involved with the environment, artifacts, etc.

Interpretation is thought to serve several purposes in recreational agencies. These purposes have changed with time. For example, Tilden (1977) suggested that interpretation's primary purpose is to provoke the visitor. Sharpe (1982) has suggested three other purposes for interpretation: 1) to assist the visitor in developing a better awareness, appreciation and understanding of the resource; 2) to accomplish management goals; 3) to promote public understanding of the agencies' goals and objectives. Propst and Roggenbuck (1981) have proposed purposes for Corps of Engineers interpretation at recreational sites such as to "inform and educate the public with regard to the purposes and concept of operation of the water project and the historical and natural features of the area." Indiana Dunes National Lakeshore's (INDU) Statement of Interpretation lists other similar, but unique, purposes.

Recently, however, park administrators and managers have questioned whether interpretation is effectively serving these purposes. For example, they have wondered if interpretive services are cost effective, if they really add to visitor experiences or aid in managing visitor behavior. The "art" of interpretation is being viewed from a more scientific and quantitative perspective. Because the purposes of interpretation are being questioned, more and more agencies and researchers (Cooperative Extension Service, 1979; National Park Service, 1983; Corps of Engineers, 1983; Wagar, 1972; Putney and Wagar, 1973; Roggenbuck, 1979; Lime,

1979; Nielson, 1980) are turning to evaluation in order to analyze and empirically document whether interpretation is meeting its intended purposes.

Although several authors (St. Clair, 1972; Putney and Wagar, 1973; Wagar, 1976; Wagar, et al, 1976; Roggenbuck, 1979; Lime, 1979; Risk and English, 1983; Hodgson and Fritschen, 1983; Moses, et al, 1977; Knudson and Morfoot, 1979) have stressed the importance of evaluation, "little systematic evaluation of the effectiveness of programs...has (actually) taken place," (ERIC, 1981). Additional research which documents the effectiveness of interpretive programs is necessary (Morfoot and Blake, 1979).

In addition, the methods which are used for evaluation of interpretation are not well developed. "The field of environmental education and interpretation has traditionally been plagued with the problem of needing to evaluate the effectiveness of programs...while lacking the techniques necessary to show that positive effects are occurring," (Elliot, 1979). Because interest in evaluation has recently heightened, numerous methods for evaluation have been developed or borrowed from other fields of study (e.g., recording quizboard, survey, observation of audience attention, etc.). These methods are being used to determine if interpretation is effectively serving its purpose. Interpretive evaluation studies, however, have not examined what each of these techniques really measures. Instead, researchers have drawn inferences about these methods based upon their application in other fields of study.

The purpose of this study was to add to the body of knowledge on evaluation of interpretive effectiveness. Part one was an experiment designed to evaluate Indiana Dunes National Lakeshore's dune succession programs for school groups. Participants' knowledge, attitudes and behavior were measured before and after exposing them to interpretive programs (treatment) and compared with a control group's knowledge, attitudes and behavior responses. Knowledge was measured with game-type quizzes. Attitudes and minimal demographic information were solicited with a short questionnaire. Finally, behavior was measured before the treatment with a "story" questionnaire and after using observation. The second part of this study examined four evaluation methods in order to see if these methods could be used singly or in a certain combination to measure attitudes, behavior and knowledge.

PROBLEM JUSTIFICATION

The Need for and Overview of Evaluation

Evaluative research (or evaluation) is "a method for studying effectiveness" of public programs (Suchman, 1967), and evaluation researchers examine the degree to which pre-specified objectives are met (Poister, 1978; Lauffer, 1978; Rossi, 1972). Evaluation is a tool to determine whether new discoveries and knowledge are being successfully applied. "All social institutions" such as education, public health, social casework juvenile delinquency, public policy, agricultural development, etc., must evaluate their programs to prove "their legitimacy and effectiveness in order to justify society's continued support" (Suchman, 1967).

The need for evaluation in all areas of public service has increased because three significant changes have taken place (Suchman, 1967). First, social problems have become more diverse and complex. The service institutions which deal with these problems are being challenged. Second, agencies are taking a broader approach to solving problems and more highly trained practitioners are increasingly needing to prove their worth. Finally, a better educated and more sophisticated public is less willing to accept community services on face value alone. The "obvious" benefits of social programs are no longer obvious when providing these benefits results in larger budgets and in turn increased tax payments.

The need for evaluation is further evidenced by the

existence of several journals such as Evaluation Quarterly, Evaluation Studies and Review Annual, Evaluation Review and Evaluation and Program Planning which report on recent studies in evaluation.

Although evaluative research is greatly needed, too few evaluations have been conducted and most that have taken place are of poor quality (Suchman, 1967; Theobald, 1979). This situation prevails for several reasons. The methodology and theory to support evaluative research is not well defined. Evaluation is inherently subjective as well as complex (Kidder, 1981; Rossi, 1972). For instance, because of its applied nature, evaluation research must always be conducted in light of political, social and/or economic ramifications of the results. Evaluation research is difficult to administer. The hypotheses are largely stated as objectives (Suchman, 1967), however, the objectives may be based on unsound assumptions and/or they may be unclearly defined (Suchman, 1967; Poister, 1978).

Evaluation thrusts have changed through time. Initially evaluations were based on authorities' view of what was good for their clientele (Suchman, 1967). Although some statistical evaluation indices were developed in the early 1700's, service was more important than evaluation (Suchman, 1967). After World War II, more demand for critical self-appraisal began. The purpose of this "new" evaluation was to increase standardization in programs and provide incentives for meeting these standards, yet standards for the programs were

necessarily arbitrarily set (Suchman, 1967). Now evaluative research tries to measure the degree to which pre-set objectives are met (Poister, 1978; Suchman, 1967; Kidder, 1981).

Evaluation in Parks and Recreation

The progress of evaluation in parks and recreation is similar to other fields. Events such as the tax revolt have caused increased pressure for evaluation of park and recreation services (Lamb and Crompton, 1980; Theobald, 1979). Like other public services, little evaluation has taken place (Chubb and Chubb, 1981; Theobald, 1979). Early efforts focused on the administrative viewpoint of effectiveness and efficiency (Chubb and Chubb, 1981; Lamb and Crompton, 1980; Theobald, 1979). For example, the number of users, expert judgement, range of programs, numbers of acres of land, etc., were used to assess park and recreation agency worth (Lamb and Crompton, 1980; Theobald, 1979).

Attendance has been the most frequent criterion used to measure service success because it appears to offer several benefits (Lamb and Crompton, 1980). It is reasonably easy to use. It has apparent objectivity. It seems politically non-controversial. It seems logical given that other decision making criteria are lacking. For instance, it is difficult to measure the quantitative value of a beautiful vista. Attendance as a measure by itself, however, has many shortcomings (Lamb and Crompton, 1980). It is not necessarily an accurate measure. If, for example, practitioners know their budget is being assessed from attendance figures, traffic

counter loading factors can be adjusted. New programs cannot be evaluated using this method. Equally important services may be designated for smaller customer groups and/or the social or ecological carrying capacity may limit the program's attendance. Substitutes may be available which cause considerable decline in attendance. Finally, a small group of highly committed customers may in fact provide more support for the program than large numbers of people who are minimally committed.

Standards have also been traditionally used as a measure of effectiveness (Theobald, 1979). Standards include the number of facilities necessary per 1000 people, the numbers of acres of parkland that a given community should have, etc. As standards are arbitrarily set, are based on historical tradition, and do not take users' needs into account, (Theobald, 1979) they are not a true measure of effectiveness. They should serve as park and recreation agency baseline operating guidelines and not as effectiveness measures.

Only recently have park and recreation personnel begun to examine their services in terms of their impact on the clientele. Additional work still needs to be done. Practitioners need a clear understanding of their program's intended impact on customers (Lamb and Crompton, 1980). They should write program objectives that indicate customer benefits based upon the program's desired impact (Lamb and Crompton, 1980; Theobald, 1979). Finally, well designed program evaluation must be implemented in a comprehensive

manner (Theobald, 1979).

Historical Progression of Interpretive Evaluation

In the past, the art of interpretation was always viewed as an essential element of recreational services (Sharpe, 1982). Recently, however, the reasons for providing interpretation have been questioned. For instance, budget cuts have caused administrators to question whether interpretation is cost efficient (Bernard, 1977; Nielson, 1980). Others have wondered if interpretation is benefiting the visitor. For example, Dawson and Roggenbuck (1979) found that the activities that visitors liked were negatively correlated to the activities interpreters liked and types of programs which interpreters liked to give. Obviously, one must question why interpretation is still provided. As a result of these researchers and others inquiries and criticisms, practitioners realized that they needed to document the impacts of interpretation.

Interpreters conducted the initial evaluation studies (if evaluation took place at all) (McDonough, 1981). The first studies took an internal perspective, and focused on the interpreter and the interpretive message/technique (McDonough, 1981; Fritschen, 1980). For example, messages were evaluated according to whether they held the audience's attention, met Tilden's principles, or were of suitable length (Propst and Roggenbuck, 1981). An interpreter was evaluated in terms of the "interpretive attributes", such as presentation organization, verbal skills and general appearance, that he/she possessed. Methods such as peer evaluation,

content analysis, audit by an expert and review by a panel of outside experts were used to evaluate the interpreter, and the message. (Table 1 details the aspects of the message and the interpreter characteristics that were evaluated.) Yet these early "internal-looking" evaluations were only indicators; they failed as precise measures of interpretive effectiveness (Propst and Roggenbuck, 1981).

Researchers recognized the need to redefine interpretive effectiveness from an audience perspective (Wagar, 1976; Wagar and Washburne, 1972; Mahaffey, 1969). To conduct evaluation from an audience perspective, interpreters set objectives which specified the program's impact on the audience. Achieving program objectives was then used as a basis for evaluation. For example, Wagar (1980) and Hodgson (1977) have stressed clear, measurable objectives are necessary for any interpretive program particularly if the program is to be evaluated. Also, Propst and Roggenbuck (1981) stressed that program evaluation objectives are needed for evaluation of Corps of Engineers interpretive programs. Nielson (1980) used objectives as a basis for evaluating her study at Grand Teton National Park. Program effectiveness in Nielson's study was defined as the extent to which pre-set objectives have been met. Nielson's approach to evaluating interpretive effectiveness parallels evaluation techniques used in other fields of study (Poister, 1978; Suchman, 1967).

Gradually "interpreter perspective" and "visitor perspective" effectiveness criteria have been integrated to

Table 1. Characteristics of the Interpretive Message and Interpreter Which are Used as Evaluation Criteria*

Message Characteristics	Definition	Implication for Using the Principle
reinforces content	uses repetition; says the same thing in many different ways	easier for audience to remember
flows from point to point	message moves logically from one point to the next with smooth transitions	helps organize information for audience so it is easy to understand
is brief	short	does not overwhelm audience
is grammatically correct		self-explanatory
uses descriptive, active, informal, emotional and/or concrete language and no jargon	self-explanatory	uses words/language that the audience can understand
uses information that people like		people like bizzare, unusual, dynamic, humorous information, information about wildlife and that shows relationships to humans
is simple		easier for audience to understand
uses questions	plan a variety of question types, e.g., close ended, open ended, etc.	encourages audience interaction and participation
is illustrated with analogies		helps the audience understand better
has unity	information is logically tied together	easier for people to understand the main concept (or theme) rather than isolated facts (also related to flow)

Table 1 (cont'd.)

Message Characteristics	Definition	Implication for Using the Principle
uses "attention getters"	devices, example surprise, to capture audience's attention	self-explanatory
uses visual, auditory aids		can enhance the program
Interpreter Characteristics		
presents a holistic perspective and is well organized		
communicates clearly, adequately loudly, has good verbal and nonverbal skills		
is enthusiastic		
has a pleasant appearance and demeanor		
is credible and knowledgeable of subject area		
is gregarious		
is self confident		
is creative		
is articulate		
has a sense of humor		
uses Tilden's principles reveal, relate, provoke example - surprise to reveal ideas		

*The characteristics listed above necessarily overlap. For instance, an interpreter reinforces the message but the message (i.e., written word, etc.) can also "contain" reinforcement.

define the reasons for conducting evaluation of interpretive effectiveness. Propst and Roggenbuck (1981) have listed four reasons for carrying out evaluation at Corps reservoirs: 1) to solicit greater agency support; 2) to assess appropriateness of objectives; 3) to determine whether objectives are being met in a cost efficient manner; and 4) to know reasons for program effectiveness. Iozzi and Cheu (1979) and Hodgson and Fritschen (1983) have suggested the reason to evaluate interpretive services is to determine if a desired behavior change has resulted. Sharpe and Gensler (1977) have said that evaluation is necessary to see if management objectives are being met. These reasons for evaluation have led researchers to question what "effectiveness" really means, i.e., how it is operationally defined and what techniques or methods can be used to measure it.

Interpretive Evaluation Studies

Interpretive effectiveness has been defined in many ways including knowledge recall, satisfaction (attitude), cognitive differentiation, and behavior change. Hammitt (1981) hypothesized that recall familiarity and visual preference were interpretive effectiveness measures. Knudson and Morfoot (1979) defined effectiveness as cost-efficiency and conducted an evaluation study that measured cost effectiveness in terms of labor and cash costs of providing interpretive services per visitor contact hour. Putney and Wagar (1973b) also measured effectiveness as cost per visitor contact.

Many studies have operationally defined effectiveness as knowledge recall. Wagar (1972), St. Clair (1972), Mahaffey (1969), Monroe (1983), Shiner and Shafer (1975) and Washburne and Wagar (1972) all performed studies where knowledge retention was measured. Lee and Uzzell (1980) have stated that interpretive effectiveness is more than knowledge retention. In order to approach effectiveness in terms of knowledge in a more sophisticated way, their studies have measured cognitive differentiation, using Piaget's theory of cognitive differentiation as a theoretical framework. Cognitive or schemata differentiation theory proposes that as one's knowledge about a subject increases, his/her mental schemata become more differentiated or organized. Lee and Uzzell (1980) found that visitors schemata on farms was more differentiated after they went to an interpretive program than before.

Effectiveness has also been defined as a change in attitudes. Ham and Shew (1979) conducted a study that examined the activities that both visitors and interpreters liked. Crompton and Sellar (1980) compiled a review of several studies that have looked at attitude changes.

Still others have suggested that interpretive effectiveness means a change in behavior. Fritschen (1983), Dick, et al (1975); Shiner and Shafer (1975) have described studies that measured behavioral changes. Table 2 details interpretive effectiveness studies that have been conducted and the operational definitions of effectiveness that have been measured.

Table 2. Type of Interpretive "Effectiveness" Measured and Method Used to Measure Effectiveness

Studies that used one definition of effectiveness		
Study	Definition of Effectiveness	Method Employed
Knudson and Morfoot (1979)	Cost effectiveness	Record of: interpreter time program attendance program time
St. Clair (1972)	Knowledge retention	Questionnaire
Hammitt (1980)	Recall familiarity and visual preference	Photo questionnaire
Ham and Shew (1979)	Activities that visitors liked	Questionnaire
Roggenbuck (1983) (in Fritschen 1983)	Behavioral change	Observation - behavior traces actual behavior
Dick, Myklestad and Wagar (1975)	Audience attention (behavior)	Observation
Shiner and Shafer (1975)	Audience exhibit viewing time (behavior)	Observation
Tai (1981)	Trail users' perceptions	Observation Interview On-site and post-visit questionnaires
LaHart and Bailey (1975)	Amount of litter left on trail	Experiment, observation Behavior and behavior traces
Lipman and Hodgson (1978)	Behavior after cave program	Record of number questions asked
Gustke and Hodgson	Enjoyment as measured by rate of travel on trail	Observation

Table 2 (cont'd.)

Studies that used more than one definition of effectiveness		
Study	Definition of Effectiveness	Method Employed
Crompton and Sellar (no date)	Attitudes	Variety of methods - usually questionnaire or survey
Lee and Uzzell (1980)	Cognitive differentiation, attitudes	Questionnaire/ survey
Mahaffey (1969)	Knowledge retention, interest	Questionnaire
Cherem (1983) (in Fritschen 1983)	Attitude change, behavior change	Report form for accidents
Freed (1983)	Knowledge change, Behavior change	Personal interviews Observation
Dartington Amenity Research Trust (1978)	Knowledge comprehension, Factual recall, Attitudes	Interview Questionnaire Observation
Peart (1979)	Attitudes, Knowledge gained, Behavior	Questionnaire Knowledge tests Unobtrusive observation
Nielsen (1980)	Increased knowledge, increased score of, attitudes	Questionnaire of attitudes and beliefs, posttest only control group
Kuehner (1972)	Behavior, Knowledge gained, and knowledge recall	Observation Questionnaire
Fritschen (1980)	Perceived knowledge gained, Attitude (enjoyment)	Self-administered questionnaire

As time progressed, more and more studies have combined more than one definition of effectiveness in order to get a "better handle" on what effectiveness really is. It is really necessary to use more than one measure to determine effectiveness (Propst and Roggenbuck, 1981). Table 2 also lists studies that have employed a combination of these definitions.

Interpretive Evaluation Methods

Numerous methods have been used to carry out evaluation of interpretive effectiveness. These techniques have been employed to measure the different operational definitions of interpretive effectiveness described above (such as factual recall, levels of cognitive complex processes, existing orientations and changes in attitudes and positive growth and change in behavior (Iozzi and Cheu, 1979). Most of these evaluation methods have been borrowed from other fields and a great deal of literature is available on them (Lange, 1980; Webb, et al, 1970; Campbell and Stanley, 1963; Kardiz, 1979; Bochner, 1979; Babbie, 1973; Poister, 1978; Babbie, 1983; Sommer and Sommer, 1980; Crano and Brewer, 1973).

These evaluation techniques have been modified and applied to the interpretive setting. Interpretive researchers have discussed how to use these methods in interpretive evaluation (Wagar, 1976; Blahna and Roggenbuck, 1979; Risk and English, 1983; Lime, 1979; Wagar, 1972; Moeller, 1980; Hodgson and Fritschen, 1983; Burrus-Bammel and Bammel, 1979). Propst and Roggenbuck (1981) have listed thirteen techniques to use in interpretive evaluation. The recording quizboard

is an efficient tool for evaluation (Wagar, 1972). Roggenbuck (1979), Lime (1979) and Risk and English (1983) have asserted that observation is a suitable method for evaluating interpretive effectiveness. Moeller (1980) suggests that the informal interview is a valuable tool for evaluation. Table 3 details the types of methods that interpretive researchers have recommended to evaluate interpretive effectiveness.

The above discussion shows that interpretive effectiveness studies have examined different definitions of effectiveness using different measurement methods. However, these studies have not documented which definition of effectiveness most suitably defines what effectiveness is or which method of measuring effectiveness best measures which type of definition.

Table 3. Recommended Evaluation Techniques to Use in Interpretive Effectiveness Research

Author(s)	Type of Method Recommended
Risk and English (1983)	Observation
Lime (1979)	Observation
Wagar (1972, 1976)	Recording quizboard - self-testing device
Burrus-Bammel, Bammel (1979)	Systematic unobtrusive observation
Hodgson and Fritschen (1983)	Experiment, quasi-experiment
Gramann and Field (1977)	Interpretive activity inventory (for observation)
Roggenbuck (1979)	Field experiment
Burch (1964)	Observation
Moeller, et al (1980)	Informal interview
Meis and Davis (1981)	Time lapse filming and video taping
Blahna and Roggenbuck (1979)	Recording quizboard
Cherem (1978)	Visitor employed photography

STUDY OBJECTIVES AND HYPOTHESES

This study was designed to accomplish two tasks:

1) To test the effectiveness of several site specific interpretive programs.

   H_o: There will be no difference between prestipulated external criteria (program objectives) and participants' responses for knowledge, attitudes and behavior.

   H_o: There will be no difference for participants' percent of correct answers for each level of schematic differentiation about dunes (both pretest and posttest).

   H_o: Participants' knowledge pretest and posttest responses will not change.

   H_o: There will be no difference between participants' attitude pretest and posttest responses.

   H_o: Participants' pretest behavior (intention) will be the same as posttest behavior.

2) To test if four different methods for evaluating interpretive effectiveness measure the same trait.

   H_o: There will be no difference in measuring environmental behavior using a questionnaire, observation, behavior intention story, or game-type devices (crossword puzzle and matching game).

METHODS

Since studies comparing various measures of interpretive effectiveness were not commonly found in the (interpretation) literature, and the methods used to evaluate interpretive effectiveness have been borrowed from another disciplines a model based on other fields of study was developed to meet the study objectives. This chapter describes the conceptual background for the study design, the study design itself, the instruments which were developed, the location of the study, study population and the details of how the study was carried out at INDU.

Conceptual Background

Study Objective 1: Evaluation of program effectiveness

Researchers have suggested numerous methods to use to evaluate the effectiveness of interpretive programs (Lime, 1979; Wagar, 1973; Fritschen, 1980; Dick, et al, 1973; Tables 1 and 2). These methods fall into two broad categories: 1) evaluation of the interpreter and his/her message; and 2) evaluation of audience response (Propst and Roggenbuck, 1981). Methods of measuring audience response fall into three major classes: 1) quizboards and other game-type devices; 2) behavioral and behavioral-trace observation; and 3) questionnaires and interviews. These three types of methods seem to best measure particular kinds of human behavior (Wagar, 1972; Robinson and Shaver, 1973; Fritschen, 1980; Sommer and Sommer, 1980; Loether and McTavish, 1974).

There are basically three categories of human behavior:

cognitive or mental behavior, affective behavior, and overt or psychomotor behavior (Peyton, 1982; Fishbein and Azjen, 1975). The first type of behavior, cognitive behavior, has been defined as knowledge recall and more recently as schema differentiation, in interpretive research (see problem justification for elaboration on schema differentiation). Questionnaires and quizboard games have been most commonly used to measure cognitive behavior in interpretive research (and other fields of study) (Wagar, 1972; Robinson and Shaver, 1973; Lee and Uzzell, 1980). A second type of behavior, overt or psychomotor action, has been measured primarily by observation (Risk and English, 1983; Lime, 1979; Hodgson and Fritschen, 1983; Sommer and Sommer, 1980; Loether and McTavish, 1974; Quick and Davis, 1979). Clark (1977) has stated that observation most precisely measures overt behavior. Also, the State of the Art on Methodology for Studying Environmental Perceptions, Belief, Attitudes and Values and State of the Art on Utilizing Perception, Attitude and Opinion Research (1974) suggested that questionnaires and similar tools which ask behavioral intention questions are better measures of behavior than attitudinal questions. Some research has suggested that people's behavior may be measured by asking them about their behavior (Haire, 1950; Lansing and Heyns, 1959). Affective behavior, the third kind of behavior, which includes attitudinal and emotive behavior, has normally been measured through questionnaires and interviews (Clark, 1977; Loether and McTavish, 1974; Crandall and Slivken, 1980).

Previous experience has indicated that no one evaluation technique is adequate to measure a program's effectiveness (State of the Art...Research, 1974; Propst and Roggenbuck, 1981). Therefore, four methods using four different types of instruments (paper and pencil games, questionnaires, behavior intention stories and observation) were designed to measure changes in three kinds of behavior (cognitive, affective, psychomotor) which might result from exposure to an INDU dune succession program.

See "Study Design" and "Instrument Development" sections for details.

Study Objective 2: Evaluation of interpretive methods

In the area of interpretation, researchers have inferred what methods are most suitable for measuring the operational definitions of effectiveness. No one is empirically sure which methods are valid measures of the different operational definitions of effectiveness (e.g., knowledge recall, attitude shift, behavior change, etc.). Yet, "every method's validity has to be empirically tested in the different situations where it is to be used" (Magnusson, 1967). In other words, do the methods that are used in interpretive evaluation research measure the traits that they are intended to measure?

In other fields, measuring concepts or traits using several methods and then looking for convergence (i.e., correlation) among these methods has been an excellent way to validate the concepts and the methods (Kidder, 1981; State of the Art...Research, 1974).

In fact, Campbell and Fiske (1959) have developed a validity test which uses correlations to test whether two or more methods measure the same trait. This technique, called the multitrait-multimethod matrix, has been used in psychological testing and similar research (Kidder, 1981; Magnusson, 1967; McGaw and Watson, 1976; Anastasia, 1968). To use this test, one or more traits must be analyzed by two or more methods that are as different as possible. Also, to account for individual differences, each participant must be tested with all the methods.

This test is appropriate to meet this study objective. Interpretive effectiveness can be defined as a change in behavior as a result of exposure to an effective program. In fact, Iozzi and Cheu (1979) have commented that some consider this definition of evaluation the "epitome of environmental education (and interpretation) achievement." In this study, "environmental behavior"was the trait that was examined and it was operationally defined as cognitive behavior (knowledge), affective behavior (attitudes) and psychomotor behavior (behavior). This operationalization is logical for several reasons. Changes in knowledge, attitudes and physical actions are expressed behaviorally. For example, people talk (a behavior) to express knowledge and attitudes. Knowledge, attitudes and actions are interrelated in any particular individual (Schafer and Tait, 1981) (Figure 1). The environmental behavior trait was measured using four different methods (paper and pencil games,

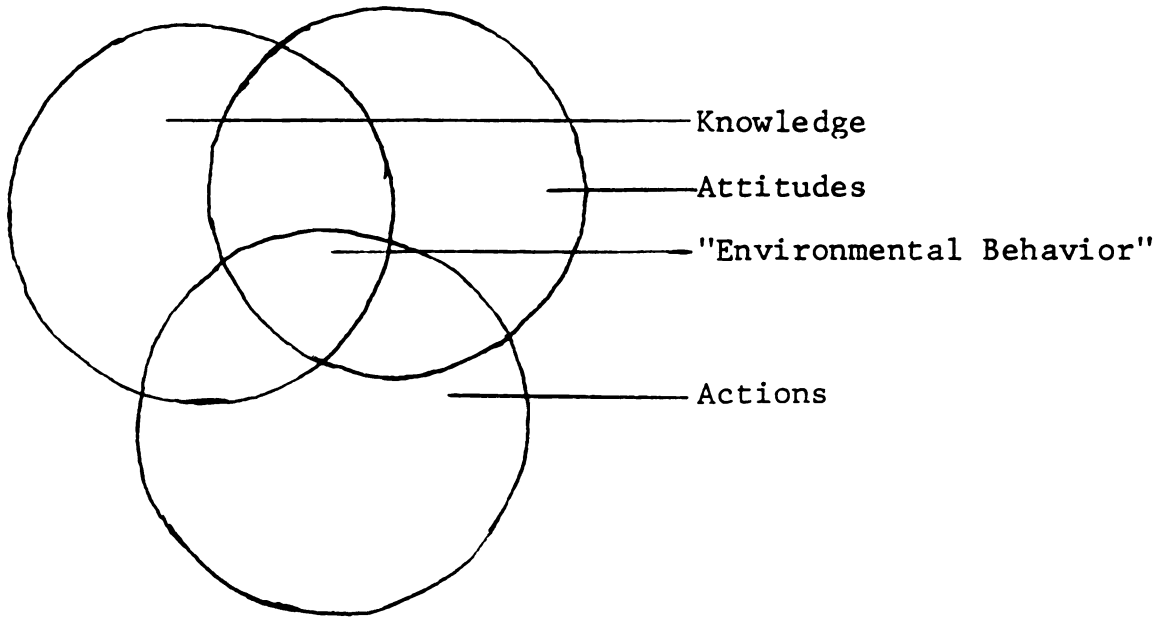


Figure 1. A model of the interrelatedness of attitudes, knowledge and actions.

questionnaires, behavior intention stories and observation). The scores from these methods were correlated to determine if environmental behavior could be measured using one or some combination of the four methods.

Study Site

A U.S.D.I. National Park Unit, Indiana Dunes National Lakeshore, is a"patchwork" land area which lies between Gary and Michigan City, Indiana (Figure 2). The study took place at West Beach, the western most area of the park. It was chosen for several reasons. The area is very popular and the impact created by such large numbers of visitors has caused numerous management concerns. Human impact such as littering, erosion, houses, sandmining and damaged vegetation is easy to see. The site also has many natural secrets. The processes of succession are easily experienced as one walks from the beach to the parking lot on dune trails. The combination of a natural environment and an environment damaged by humans made this site perfect for the study because both of these aspects were woven into the program and then examined in the evaluation.

Study Design

A pretest/posttest control group design was chosen because it offered several advantages for this particular study. An experimental design is recommended when evaluation research is conducted (Houston, 1972; Kidder, 1981; Suchman, 1967; Anderson and Ball, 1978). Hodgson and Fritschen (1983) have suggested two experimental designs to

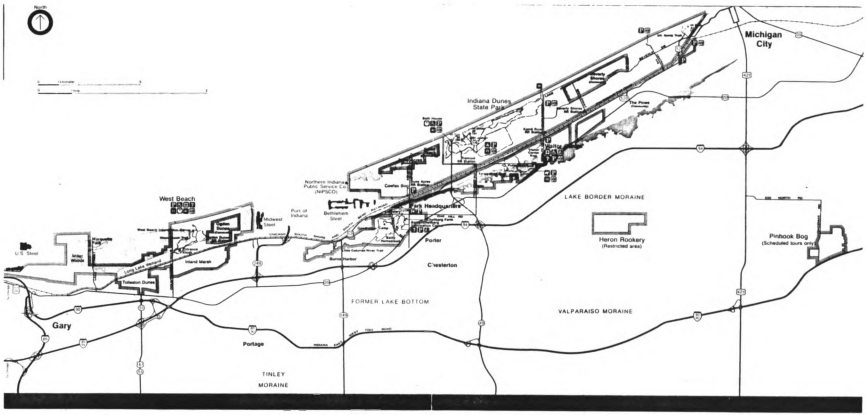


Figure 2. Indiana Dunes National Lakeshore

to evaluate interpretive services, including the pretest/ posttest control group design.

In previous interpretive effectiveness studies, many variables were not controlled (Fritschen, 1980). The experiment helps control for some of the variables. The presence of a control group enabled the researcher to attribute any effects of the treatment to the treatment and not other factors and see if any changes may have resulted from the treatment. An experiment does not require the large numbers of participants which are necessary for survey and other modes of research. Both pretest and posttest results can be compared to prestipulated external conditions (program objectives - see "Instrument Development" section). This design can be described as follows:

$R_1 O_1 X \; O_2$

$R_2 O_1 \; O_2$

R_1 is the experimental group

R_2 is the control group

O_1 is the pretest

O_2 is the posttest

X is the treatment

Four instruments were used as pretest and posttest measures. Participants' pre-treatment and post-treatment knowledge was measured using pencil and paper games. Attitudes were measured using a short questionnaire for both the pretest and posttest. Pre-treatment behavior was measured using a behavior intention story and observation was used to measure post-treatment behavior (Table 4).

Table 4. Number of Individuals in the Control and Experimental Groups and the Instruments That They Filled Out.

School Name	Number of Individuals Study Objective 1	Instruments Completed for Study Objective 1	Number of Individuals Study Objective 2	Instruments Completed for Study Objective 2
		Experimental Group		
Bennett	17	crossword pretest crossword posttest questionnaire pretest questionnaire posttest behavior intention story	6	crossword pretest crossword posttest questionnaire pretest questionnaire posttest behavior intention story observation
Pieffer	33	crossword pretest crosword posttest questionnaire pretest questionnaire posttest behavior intention story	12	crossword pretest crossword posttest questionnaire pretest questionnaire posttest behavior intention story observation
Edgewood	21	matching pretest matching posttest questionnaire pretest questionnaire posttest behavior intention story	6	matching pretest matching posttest questionnaire pretest questionnaire posttest behavior intention story observation
		Control Group		
Pieffer	23	crossword pretest crossword posttest questionnaire pretest questionnaire posttest behavior intention story	6	crossword pretest crossword posttest questionnaire pretest questionnaire posttest behavior intention story observation

To meet objective one, group size and mode of interpretation were systematically varied (independent variable). In other words, the type of program (or treatment) was the independent variable (Table 5). The effectiveness of the program, measured in terms of behavior (cognitive, affective and overt) change, was the dependent variable. The variables for meeting objective 2 were different. The method of evaluation (games, questionnaires, behavior intention story, or observation) was systematically varied (i.e., was the independent variable). The validity correlation (the measures of what the method tests) that resulted from varying the methods was the dependent variable.

Park personnel terminated data collection in the middle of the study. Therefore only very few individuals were exposed to each treatment. In most cases individuals from a particular school were given the same treatment and hence could not be compared with others in the event that their scores and responses were unusually high or low. In addition, the group size was reasonably similar (20-30 people) rather being distinctly small or large. Data from only one type of treatment were collected for the control group. As a result of the above situation, all treatment groups were "lumped" together and treated as an experimental group (Table 5). This group was then compared with a single control group.

Table 5. School Groups Assigned to Experimental and Control Treatments: Original and Final Designs

Original Design

	Experimental Group				Control Group				
	Mode of Interpretation**								
	Lecture		Activity		Lecture		Activity		Row
Group Size†	(School name, n_1*n_2*)		(School name, n_1, n_2)		(School name, n_1, n_2)		(School name, n_1 n_2)		Total
Small	Bennett	17 (6)	Peiffer	15 (6)	Springfield	15 (6)	Lincoln	15 (6)	
	Springfield	15 (6)	Lincoln	15 (6)					
	Scout Troop 199	15 (6)							
		47(18)		30(12)		15 (6)		15 (6)	107(42
Large	Edgewood	27 (6)	Peiffer	30 (6)	Skiles	33 (6)	Peiffer	30 (6)	
	Skiles	32 (6)	Skiles	32 (6)			Springfield	30 (6)	
	Lincoln	30 (6)	Skiles	33 (6)					
		89(18)		95(18)		33 (6)		60(12)	277(54
Column Total		136(36)		125(30)		48(12)		75(18)	384(96

Final Design

Experimental Group		Control Group	
(School name, n_1*, n_2*)		(School name, n_1*, n_2*)	
Bennett	17 (6)	Peiffer	30 (6)
Peiffer	40 (6)		
Edgewood	27 (6)		
Column Total	84(18)		30 (6) 134(24)

*n_1 = number of individuals for study objective 1 (evaluation of program effectiveness)
n_2 = number of individuals for study objective 2 (evaluation of interpretive methods)
**mode of interpretation: activity program - incorporates activities, encourages audience participation;
lecture program - does not incorporate activities, does not encourage audience participation;
not used in final design
†small: approximately 15 people; large: approximately 30 people; not used in final design

Study Sample

The study population consisted of organized groups of children in fourth to sixth grades. INDU staff sent an "invitation to participate" and general information about West Beach (Appendix A) to over 1200 organized groups in Indiana and Illinois.

Out of 1200 organized groups, nine school groups and two scout groups signed up for the study program (Table 5). Originally, 484 individuals made up the study sample. However, some cancellations occurred and data collection was terminated, resulting in a total sample size of 112 individuals from three schools. All these individuals filled out the pretest (questionnaire, game and story) instruments in order to meet study objective 1 (Table 4).

To meet study objective 2, it was necessary for each individual to be exposed to all four evaluation methods. In the classroom, where pretesting took place, it was possible for each individual (n=112) to fill out all three instruments. However, because an observation strategy instead of a behavior intention story was employed for the posttest, it was not possible to observe all individuals in all groups. Therefore, six individuals from each group were randomly selected before they arrived at the park to be observed. They, as well as all others, completed the posttest knowledge and attitude instruments. The individuals who were observed were identified with various colored tags. All other group members were given a single color tag. The

colors were tied into the program with a discussion on fall leaf colors. The total number of individuals to meet study objective 2 was 120 before the school and scout groups cancelled and 27 after the study was curtailed. A complete list of the number of individuals in each school group who participated in the program is found in Table 4.

Before individuals arrived at the park, they were randomly assigned to treatments by school group. Three experimental groups (but were treated as one, see Table 5) (n=84) and one control group (n=18) were chosen. Kinds of knowledge tests (matching, or crossword puzzle) were also randomly assigned to the different groups. (Because data collection was terminated, no control group filled out the matching.)

Instrument Development

Eleven instruments were developed for this study. Instrument content was derived from the following program objectives (and program content, Appendix B) which INDU Environmental Education staff prepared.

Cognitive Behavior (Knowledge) Program Objectives

By the end of the program:

1) 90% of the participants will know that marram grass and other plants help hold the dunes in place and prevent erosion.

2) 90% of the participants will know that walking on the grass can eventually kill it and cause erosion.

3) 80% of the participants will know that plants and animals are protected and shouldn't be disturbed in national parks.

4) 70% of the participants can correctly indicate a common dune animal's and plant's habitat.

5) 70% of the participants can correctly identify one successional vegetation community

Affective Behavior (Attitudes) Program Objectives

By the end of the program:

6) 80% of the participants will want to visit the National Lakeshore again.

7) 80% of the participants will feel that it's very important to protect plants and animals in the Lakeshore.

8) 70% of the participants will report that they enjoyed the program very much.

9) 50% of the participants will report that they would rather help prevent erosion than run down a forested dune.

10) 90% of the participants will state that they feel that littering is bad.

Psychomotor Behavior (Behavior) Program Objectives

By the end of the program:

11) 80% of the participants will stay on the trail (at a given location).

12) During the program, no more than 20% of the participants will walk off the trail onto vegetated areas.

13) At least 50% of the group will properly dispose of a piece of trash which they are inconspiculously given during the program.

As much as was possible the instruments' internal (face) validity was checked. The instruments were geared to the fourth to sixth grade level. For example, words were used that were appropriate for fourth to sixth graders (Walsh, personal communication, 1983).

The instruments were also developed to work within the study environment. Index stock paper was used for those instruments (posttest games and questionnaire) that were

completed in the park. Clipboards and pencils were provided for the interpreters and the teachers so they could fill out the behavior posttest (Appendix C) and program content instruments (Appendix C) respectively.

Since the instruments were not identical, they also checked for reliability. First, the instruments were pilot tested (see pilot testing section). Second, at least two questions were used (in most cases) to measure a given objective as an increased number of similar items (up to a certain point) increase the measurements' reliability (and validity) (Garrett, 1960; Leinke and Wiersma, 1976). Third, reliability correlations were run. The pretest items of each method were tested with the posttest items for that method for the knowledge and attitude instruments.

Development of the instruments themselves followed logical and sequential steps. So each question reflected a specific objective, the operational meanings of the program objectives were determined. For example, attitude program objective seven states that "80% of the participants will feel it is very important to protect plants and animals..." A question could not be directly phrased from this objective because it is too ambiguous. Instead, a question "how do you feel about someone who throws stones at birds" more specifically measures the objective. It also asks about a specific problem found at West Branch. The questions also needed to be phrased so they were not leading. For instance, "how important do you feel it is to protect plants and

animals in the Lakeshore," is a leading question which does not measure if the objective has or has not been met. An operationalized version of the same question would be "how do you feel about protecting plants and animals in the Lakeshore?" Respondents can then answer any way they please. These steps were carried out for all instruments.

The questions were then compiled into draft instruments. Research experts such as program evaluators and educational psychologists at Michigan State University, as well as park personnel, critiqued and made suggestions about these draft instruments. From these comments, "second round" instruments were written and later pretested.

Knowledge Instruments

Three different types of pencil and paper games (cross-word puzzle, matching game, fill-in-the-blank stories (Figures 3, 4, 5, 6, Appendix C) were developed to measure cognitive behavior changes as it was not possible to use recording quizboards or similar devices in the outdoor park environment. Although several paper and pencil games (e.g., National Park Service "Habitat Game," "Salmon Game") are already available to the public, these have not been incorporated into evaluation. These instruments were designed to measure participants' levels of schematic differentiation about dunes, rather than the traditional "recall of factual information" definition. In order to analyze whether participants' mental schemata about dunes became more organized, several questions were used to measure each program objective

Name_____

Searching the Dunes

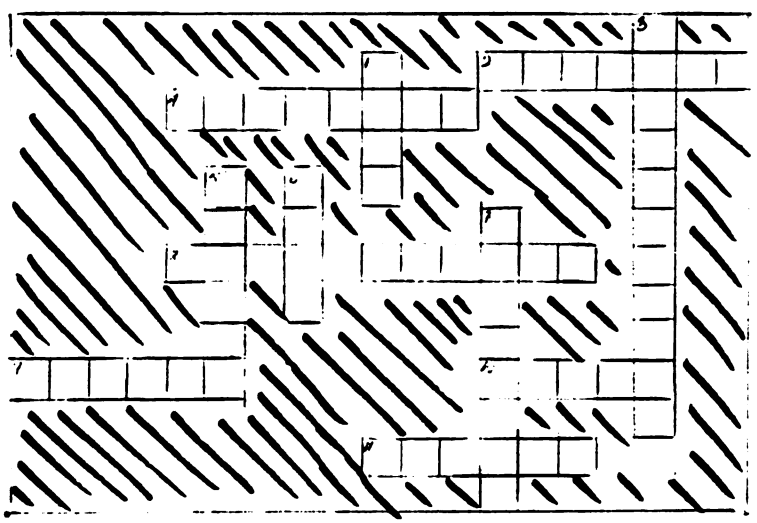
As you explore a nature park you learn many interesting things. To help you remember some of the things you have discovered, we have given you some clues. See if you can use these clues to fill out the puzzle below. HINT: Words to help you with your spelling are listed at the bottom of the page. Sometimes you may need to use the same word twice. Some extra words are also listed that you may not need to fill out the puzzle.

DOWN

1. ___blows the dunes away when no plants grow on them.
3. ___is an important plant that helps keep the dune from blowing away.
5. You may not throw ___at the animals.
6. Walking on plants will eventually ___them.
7. The area between the upper beach and the cottonwoods.

ACROSS

2. ___ are living things which move about on the dunes. They move using their legs, wings, etc.
4. This tree is the main plant found in the evergreen community.
8. An animal that lives mostly in the cottonwood community.
9. You are not permitted to take ____ home with you.
10. Hills of sand blown by the wind are ___.
11. The living things that grow in the sand.



SPELLING WORDS

marram grass(es)
hurt(s)
jackpine(s)(forest)
acorn(s)
hill(s)
wood pond(s)
stone(s)
blowout(s)
plant(s)
wind
foredune(s)
tree(s)
mammal(s)
Fowler's toad(s)
Stick(s)
rain
animal(s)
dune(s)
trash
box turtle(s)
kill(s)
hognose snake(s)

Figure 3. Knowledge Pretest: Crossword

Name_____

Searching the Dunes

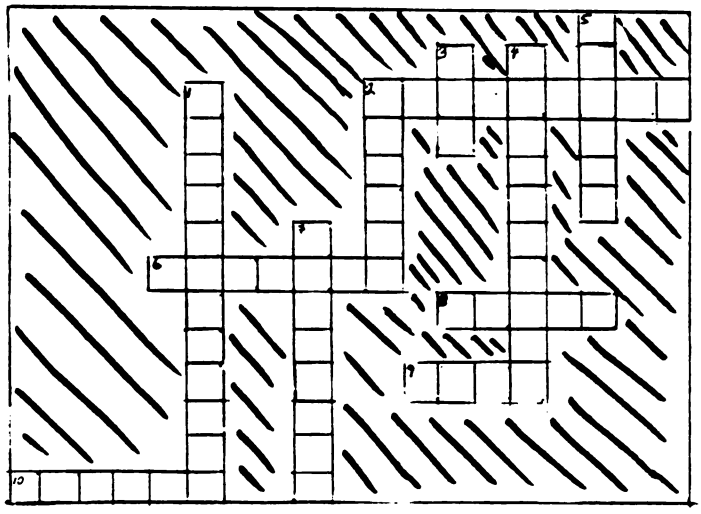
As you explore a nature park you learn many interesting things. To help you remember some of the things you have discovered, we have given you some clues. See if you can use these clues to fill out the puzzle below. HINT: Words to help you with spelling are listed below. Sometimes, you may need to use the same word twice. Some extra words are also listed that you may not need to complete the puzzle.

DOWN

1. ___are large birds that find food near Lake Michigan.
2. Green ___ have roots which burrow into the sand.
3. Plants ___when people walk on them.
4. ___trees help keep the dune in place.
5. People are not supposed to hit animals with ___.
7. The ___community is between the oak forest and the cottonwoods.

ACROSS

2. ___ ___grows in the evergreen community. HINT: The plant can cause an allergic reaction if you touch it.
6. Some ___eat plants that live on the dunes.
8. If no plants grow on the dune, the dune ___away.
9. The wind blows ___ into dunes.
10. You may not pick these things.



SPELLING WORDS

stone(s)
jack pine(s)
plant(s)
reptile(s)
cottonwood(s)
blow(s)
dirt
raccoon(s)
mallard duck(s)
tree(s)
grass(es)
sand(s)
herring gull(s)
move(s)
cry
animal(s)
poison ivy
foredune(s)
stick(s)
die(s)

Figure 4. Knowledge Posttest: Crossword

Name_____

Dune Habitat Hide & Seek

You are a scientist who has just made some discoveries about what's going on at a nature park. These important discoveries need to be recorded! Please place an "X" in the box where you think these things are going on. You may have one, two, three or four "X"s in each row. Leave the box blank if you don't know where these things are happening.

		Oak Forest	Foredune	Upper Beach	Evergreen Forest
EXAMPLE:	A squirrel lives here.	X			
	It rains in this area.	X	X	X	X
1.	The hop tree grows here.				
2.	People cannot take plants home from this area.				
3.	Plants live here.				
4.	The burrowing spider digs its home here.				
5.	Hills of sand are found here.				
6.	If no plants grow in this community, the dune blows away.				
7.	If people walk on plants in this area, they eventually kill the plants.				
8.	Marram grass hold the dune in place in this area.				
9.	Animals live in this area.				
10.	In this area, you may not throw stones at animals.				
11.	The community between the upper beach and the cottonwoods is called___.				

Figure 5. Knowledge Pretest: Matching Game

Name_____

Dune Habitat Hide & Seek

You are a scientist who has just made some discoveries about what's going on at a nature park. These important discoveries need to be recorded! Please place an "X" in the box where you think these things are going on. You may have one, two, three or four "X"s in each row. Leave the box blank if you don't know where these things are happening.

	Interdunal Pond	Oak Forest	Cottonwood Community	Fore-dune
EXAMPLE: A snake lives here.		X		
Snow falls in this area.	X	X	X	X
1. People are not supposed to hit animals with sticks here.				
2. In this community, if people walk on the plants, they eventually die.				
3. Bearberry plants grow in this community.				
4. In this area, cottonwood trees help keep the dune from blowing away.				
5. This area is made up of sand.				
6. Visitors are not allowed to pull up plants in this community.				
7. A raccoon lives in this area.				
8. Wind blows the dunes away when no plants grow on them in this area.				
9. Plants grow in this area.				
10. Animals live in this community.				
11. The community that succeeds the foredune is the ___.				

Figure 6. Knowledge Posttest: Matching Game

(Table 6). Also, the specific plants and animals used in each set of pre- and posttest items were randomly chosen and based on the program content. The content of each of the questions and how they relate to the objectives are detailed in Table 6.

In addition, spelling words were provided for the cross-word puzzle and story because many students in these grades have spelling problems. Several words which had the same number of letters were given to help eliminate guessing. The stories were based on heroes because fourth to sixth graders favor stories which are hero-centered (Lesser, 1974).

Attitude Instruments

Pretest

A short questionnaire was developed to ascertain attitudinal information from the study participants before participating in the program (Figure 7). This option was chosen over interviews because it was not possible to interview every individual.

Each question on the instrument directly reflected an attitudinal program objective (Table 7).

The questionnaire had scales so that attitudes before and after the program could be compared. The scales originally were different from each other and used words. For example, one scale ranged from lots of fun - fun - so-so - very boring, while another scale went from very annoyed - annoyed - so-so. These scales were modified after comments from park personnel and educational specialists

Table 6. Knowledge Program Objectives and the Pretest and Posttest Items That Correspond to Each Objective.

Knowledge Program Objective Number	Program Objective Content	Instrument Question Content[1]	Crossword Puzzle		Matching Game	
			Pretest Question Number(s)	Posttest Question Number(s)	Pretest Question Number(s)	Posttest Question Number(s)
1.2	know plants prevent erosion, know walking on plants kills them	no plants grow here, dunes blow away (general)	1	8	6	8
		if people walk on the plants, they die (general)	6	3	7	2
		specific plant keeps the dune from eroding (specific)	3	4	8	4
		one general question about plants	11	2 down	3	9
3	know plants/ animals are protected	one general question about protecting and disturbint plants	9	10	2	6
		one general question about protecting and disturbing animals	5	5	10	1
4	know animals have specific habitats	one general question about animals (less organized)	2	6	9	10
		one specific question about a particular animal's habitat	8	1	4	7
		one specific question about a plant's habitat	4	2 across	1	3
5	know a successional community	one question about identifying a successional community (specific)	7	7	11	11
-	-	one question about "what is a dune" (general)	10	9	5	5

[1]"General" at end or in phrase indicates that the question tested for less differentiated or general concepts. "Specific" at end or in phrase indicates that the question tested for more differentiated or specific concepts.

Name_____

We would like to know how YOU feel about visiting parks and how YOU feel about the things you do at parks. Please put an "X" on the face that best tells how you feel. There are no right or wrong answers. Thank you.

EXAMPLE: How do you feel about going to movies?

1. How do you feel about going to nature parks?

2. How do you feel when you see an empty coke can on the side of the road?

3. How do you feel when you see a left-over sandwich on the beach?

4. How do you feel when you see candy wrappers on the ground at a nature park?

5. How do you feel about people who play their radios very loudly when you are trying to relax closeby?

6. How do you feel when you see someone throw a stone at a duck?

7. How do you feel when you see someone taking plants from a nature park?

8. Suppose you could run down some sand hills at a nature park. Where would you most like to run? Choose one answer.

_____On a sand hill with plants

_____On a sand hill without plants

_____On a marked trail on the sand hill

_____Don't Know

9. Have you ever visited Indiana Dunes National Lakeshore?

_____Yes → Would you like to visit Indiana Dunes National Lakeshore again?

_____Yes

_____No

_____No → Do you like to go back to the same nature park more than once?

_____Yes

_____No

Figure 7. Attitude Pretest

Name_____

10. Who do usually go with when you visit nature parks? You may have more than one answer.

_____Parents/Family

_____School Group

_____Scout, 4-H, etc. Group

_____Friends

11. After you visit a nature park, do you ever talk to your family about your visit?

_____Yes

_____No

12. After you visit a nature park, do you ever talk to your friends about your visit?

_____Yes

_____No

13. Age:_____

14. Please circle:

Male Female

THANK YOU!

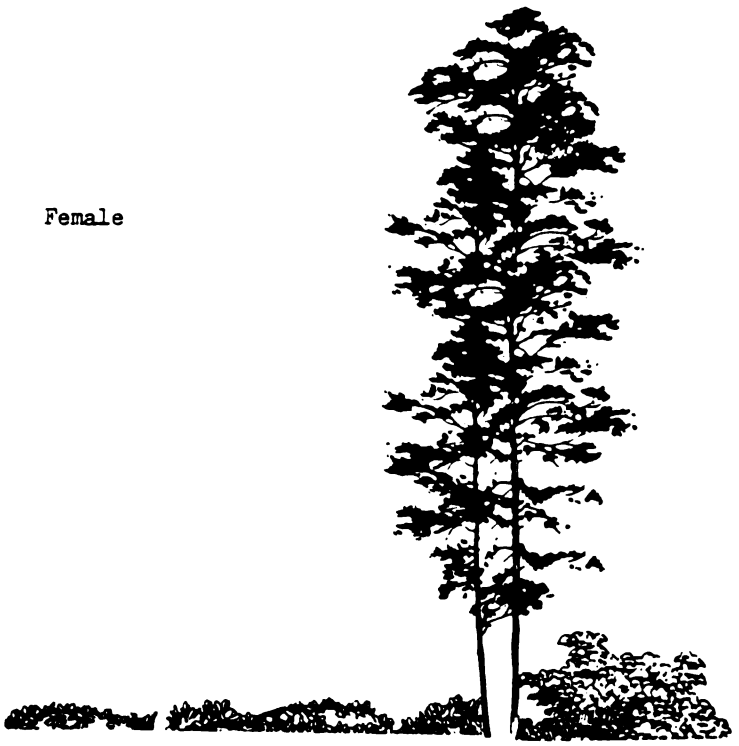


Figure 7 (cont'd.)

Table 7. Attitudinal Program Objectives and the Corresponding Pretest and Posttest Items.

Attitude Program Objective Number	Program Objective Content	Pretest Question Number(s)	Posttest Question Number(s)
6	want to visit again	1	1,2
		9	none (description only)
		9a	10
		9b	10
		10	11
		11	12
		12	13
7	feel important to protect plants/animals	5	3
		6	5
		7	4
8	enjoy the program	not applicable	2
			16
			17
9	want to prevent erosion	8	9
10	feel littering is bad	2	8
		3	7
		4	6

because they were too complex for the age group. The final scale used five smiling faces.

Posttest

A second similar questionnaire was produced to measure attitudes after the program (Figure 8). Identical instruments were not used because the pretests were usually administered only one day before participants came to the park. The questions, based on the objectives, were designed to measure the same concepts and attitudes as on the pretest (Table 7). The scales and the logic behind them were the same as the pretest attitude instrument.

(Psychomotor) Behavior Instruments

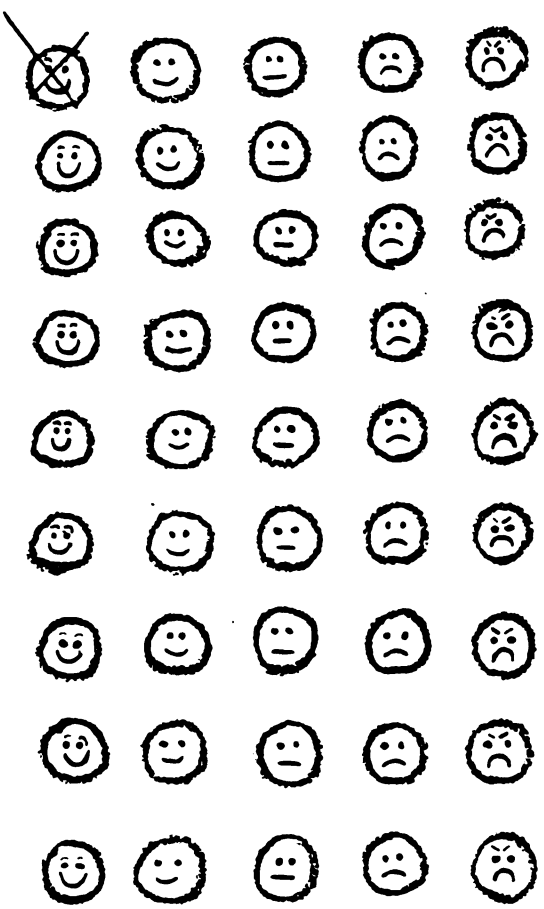
Pretest

A behavior intention story questionnaire was developed to replace observation (Figure 9) because the study individuals were not at the park prior to the program. This approach also seemed reasonable given that higher correlations have been observed between statements of behavior intention and behavior and lower correlations between general attitude questions and behavior (State of the Art... Research, 1974; Lansing and Heyns, 1959; Haire, 1950).

Two questions measured individuals' littering behavior (behavior program objective 13) and two questions measured staying on the trail behavior (behavior program objectives 11 and 12). The questions (stories) were designed to be specific and set up a realistic stiuation for participants as Kidder (1981) has suggested that specific questions are

Name_____

Today we would like to find out how you felt about your trip to Indiana Dunes National Lakeshore. Please put an "X" on the face that best tells how YOU feel. There are no right or wrong answers. Thank you.



EXAMPLE: How do you feel about climbing trees?

1. Before you came on this trip, how did you feel about coming to Indiana Dunes National Lakeshore?
2. How did you feel about the program today?
3. How do you feel about people who make lots of noise when you are trying to relax nearby?
4. How do you feel about people who collect plants at nature parks?
5. How do you feel when you see someone poke a snake with a stick?
6. How do you feel when someone leaves a crumpled piece of paper on the ground at a nature park?
7. How do you feel when you see a banana peel off the side of a trail?
8. How do you feel when you see someone throw an empty bottle on the ground?

9. If you came back to Indiana Dunes National Lakeshore, where would you most like to run? Choose one answer.

_____Down a dune covered with plants

_____Down a dune without plants

_____On the dune trails

_____Don't know

10. If you had a chance, would you like to come back to Indiana Dunes National Lakeshore?

_____No

_____Yes What would you like to do?_____

_____

Figure 8. Attitude Posttest

Name_____

11. Who would you bring if you came back to Indiana Dunes National Lakeshore? You may have more than one answer.

_____Parents/Family

_____School Group

_____Scout, 4-H, etc. Group

_____Friends

12. Will you talk to your family about your visit to Indiana Dunes National Lakeshore?

_____Yes

_____No

13. Will you talk to your friends about your visit to Indiana Dunes National Lakeshore?

_____Yes

_____No

14. Age: _____

15. Please circle: Male Female

16. What do you think is the most important thing that you learned today?

_____

_____

_____

17. What did you like best about your trip today?

_____

_____

_____

Figure 8 (cont'd.)

Name_____

A Day at the Park

We would like to know about the kinds of things that you do at parks. Please read the story and check "X" what you would do if you were in the story. There are no right or wrong answers. We would like to know what you would do.

1. You have gone to a nature park for a day of fun. One thing that you're going to do is eat lunch. After lunch, you have to take care of the wrapper from the hotdogs. What do you do with it?

   _____Put it in you pocket.
   _____Put it in the bathroom toilet.
   _____Drop it in the trashcan.
   _____Put it in the charcoal fire.
   _____Drop it on the ground.

2. You and your friends are taking a hike at a nature park that has sand hills. Grasses and bushes are on both sides of a market trail. You are walking along the trail enjoying the sights. SUDDENLY, you spot an interesting animal about three giant steps off the trail. You really want to study this animal closely. How do you get close to the animal?

   _____Run over to the animal.
   _____Stay on the trail and remain quiet.
   _____Scream to see if the animal jumps.
   _____Follow a trail that takes you to the animal from the other side of the animal.
   _____Throw cans at it to see if the animal will jump closer to you.

Figure 9. Behavior Pretest: Behavior Intention Stories

Name_____

3. A park ranger is leading your class on a walk up and down sand hills. He tells you to watch for anything interesting or exciting. When your best friend discovers a plant with large, yellow thorns, she runs off the marked trail. What do you do?

_____Run off the trail to where she is.

_____Run around a short path that takes you to the plant.

_____Look at the plant from where you are.

_____Have your friend pick the plant and bring it to the trail so everyone can see it.

_____Ask the ranger if you'll see other plants like it later on.

4. You are taking a bike trip with your scout troop to a nature park. You have taken potato chips and soft drinks for your afternoon snack. After you finish the food, what do you do with the empty chip package and the cans?

_____Throw them into the lake so they sink to the bottom.

_____Throw them in the trash barrel at the park.

_____Stuff whatever you can into your pocket.

_____Throw them into a ditch at the park.

_____Give them to another scout.

THANK YOU VERY MUCH!

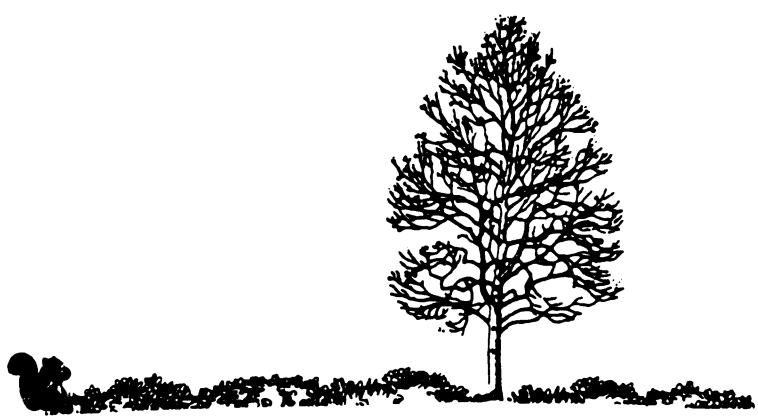


Figure 9 (cont'd.)

more likely to elicit a valid indication of future behavior than general questions.

Posttest

An observation checklist was developed to record littering and staying-on-the-trail behavior (Appendix C). The checklist, a single sheet, was set up so the researcher or interpreters could easily complete it. Question 2 was devised to record what participants did with trash (behavior program objective 13). Questions 1 and 3 (reflecting behavior program objectives 11 and 12) were designed so the interpreter or researcher could easily record whether the child stayed on or ventured off the main trail. Pretest behavior questions 1 and 4 correspond with the posttest littering behavior (question 2). Pretest behavior questions 2 and 3 correspond with question 3 on the posttest.

Content Checklist Instrument

To reduce the problem of inconsistent program content, a content checklist was generated to record what content each program covered (Appendix C). The checklist was arranged by stops (like the program content script) and was a direct result of the program content which park personnel produced. The checklist included all the plants and animals which might have been on the knowledge instruments, the biological processes, and rules and regulations, which interpreters discussed during the program.

Park Interpreter Evaluation Instrument

Since study objective 2 examined methods, it was important to gain perceptions, criticisms and suggestions from those who implemented the methods. To accomplish this goal, a questionnaire type instrument was being developed to solicit the interpreters' feelings and thoughts about the instruments (i.e., the methods) and their use. The questions included: ease of administration/use, burden on the visitor, type of detail and usefulness of information obtained, speed of feedback, cost of instrument type, guarantee against bias, difficulty in designing the instruments, comparability with other methods, amount of information provided from visitors' perspective (versus agency's viewpoint) and ease of analysis. However, because data collection was never completed, this instrument was never completed or administered.

Pilot Testing

The instruments were pilot tested at two schools in Grand Ledge, Michigan on approximately 100 students in fourth, fifth and sixth grades.

Pilot testing helped correct problems that existed in the instruments. Students were asked individually and in small groups, what troubles and problems they had with the instruments. Revisions were made according to the comments which the students gave. The final instruments reflect the revisions. For example, the story knowledge instruments were not used at the park because they were difficult and time consuming to complete in a classroom. They would have

been even more difficult to complete in a field setting.

Pilot testing also was a way to check the instruments' reliability before they were actually used. The answers were similar in all three schools suggesting repeatability over time.

Program Implementation/Data Collection

Data collection was scheduled from October 17 to November 7, 1983. Both park interpreters and M.S.U. personnel were involved in this aspect of the study. Park interpreters gave (as much as was possible) the same program over and over. M.S.U. personnel administered pre- and post-test instruments and oversaw the study.

The experimental group went forward on the West Beach trail, while the control group went in the opposite direction. When more than one experimental and/or control group came at the same time, the groups were staggered on the trail.

Park interpreters conducted two trial run programs on October 17 and 18, 1983. The first "official" program was carried out on October 19, 1983. Only posttest data were collected during this time, which were not analyzed. On October 21, 24, and 25, the complete program was implemented. Pretest and posttest data were collected for these programs. Data collection was terminated on October 25, 1983.

Administration of Instruments

Pretest for Control and Experimental Groups

The researcher distributed pretest attitude, behavior and knowledge instruments at the schools that were within a

two hour drive from the Lakeshore. At that time, study participants completed the three instruments (attitude questionnaire, behavior intention story and knowledge game) in about thirty minutes. All students were given the same amount of time to complete the instruments. The researcher collected the instruments for later analysis.

The pretest instruments (Figures 3, 7, 9) and a cover letter (Appendix D) were sent to teachers at schools that were located farther than two hours drive from the park. The teachers were requested to bring the completed forms with them to the Lakeshore the day that their program was scheduled.

Posttest for the Experimental Group

The posttest attitude and knowledge tests were administered at the end of the program at the bottom of a dune. The students took about fifteen minutes to fill out the instruments. As with the pretests, the amount of time students took to complete the instruments was kept as constant as possible.

After they had finished, the researcher gave each person a piece of candy and told them to go to the bus. The interpreter observed what the students with the different colored tags did with their candy wrappers and whether they stayed on the main trail when they went back to the bus. (Hershey kisses were used because the aluminum wrappers were easily visible if they were dropped. Non-chocolate candy (with paper wrappers) was also provided for

students who were allergic to chocolate.)

Posttest for the Control Group

The researcher administered the knowledge and attitude posttest instruments as soon as the students got off the bus, in the parking lot. She gave the students a piece of candy after they had finished their tests, and told them to meet the interpreter who was standing on the hill approximately 100 yards away.

As the children left to meet the interpreter, the researcher observed the students with the different colored tags to see what they did with their candy wrapper and if they stayed on the trail as they went to meet the interpreter.

Program Control

The teachers were given the program content checklist when they arrived at West Beach. The researcher explained the purpose of the checklist and asked the teacher to check off those facts and concepts which were covered in the program. The teacher returned the checklist to the researcher at the end of the program.

Data Analysis

Because this study involved two objectives, both of which had many variables measured at different levels, several types of statistics were used. In addition to applying the statistics to each individual question (test item), scores for each instrument were created and analyzed for both objectives one and two.

Scoring Procedure

To create the scores, a point(s) was awarded for each question if the answer was consistent(i.e., "right") with the objective. Partial credit was also given in some instances. If the answer was opposite to what the objective specified, no points were awarded to the question. Then all points were summed up to produce a total score. Table 8 lists the points that were awarded to the various questions.

The scores were calculated this way for two reasons. First, in education, grades or scores are attributed to test questions based on predetermined facts or concepts. Although the program objectives were not necessarily facts or concepts, they were predetermined conditions which participants tried to attain. Second, several books (Kidder, 1981; Magnusson, 1967; Leinke and Wiersma, 1976; Garrett, 1960) which deal with creating scores for psychological testing (such as multitrait multimethod matrix used for study objective two) give no specific information about how to create and/or calculate the scores. Hence, the author felt "grading" the instruments according to the program objectives was the most logical approach.

For each objective, the statistics are detailed below. All data were analyzed using SPSS (Statistical Package for the Social Sciences) at the M.S.U. Computer Facility. For all statistical tests, a confidence interval of 95% was used. Hence, a computed statistic of greater than .05 meant that the changes (study objective 1) and correlations

Table 8. Points Assigned to the Questions for Use in Formulating Scores.

Instrument Type	Instrument Question Number(s)	Point Value	Total Points
Knowledge			
crossword puzzle	1-11 pretest 1-10 posttest	1 point for each correct answer	11
matching game	1-11 pretest 1-11 posttest	4 points for a completely correct answer 3, 2, 1 points for a partially correct answer (for example: if two boxes were marked and four should have been checked, 2 points were assigned)	44
Attitude			
questionnaires	1- pretest 1-2 posttest 2-7 pretest 3-8 posttest 8 pretest, 9 posttest 9a, 9b; 10	1 point for 2 points for 1 point for 2 points for 1 point - "run on a marked trail" 1 point for "yes"	17 - pretest 18 - posttest
Behavior			
behavior intention story	pretest 1	2 points - "puts in pocket" 5 points - "puts in trash can" 2 points - "puts in charcoal fire"	20

Table 8 (cont'd.)

Instrument Type	Instrument Question Number(s)	Point Value	Total Points
	pretest 2	5 points -"stay on trail and remain quiet"	
		5 points -"follow a trail..."	
	pretest 3	5 points - "run around a short trail..."	
		3 points - "Look at the plant"	
		3 points - "Ask the ranger..."	
	pretest 4	5 points - "Throw in trash barrel"	
		1 point -"stuff into pocket"	
		1 point -"give to another scout"	
observation	posttest 2	5 points - "puts in trash can"	10
		2 points - "puts in pocket"	
	posttest 3	5 points - "on"	
		2 points - "on"unmarked trail	

(study objective 2) were not statistically significant.

Study Objective 1: Evaluation of Program Effectiveness

The measurement levels for knowledge, attitude and behavior variables differ. The knowledge data are interval level. The attitudinal information was measured at both ordinal and nominal levels. The behavior instruments measured nominal level data. Therefore, different statistical techniques were employed with each measurement level so important statistical assumptions were not violated. Table 9 details the statistics that were used.

Frequencies were run to get a profile of the data's distribution and to check for erroneous responses before additional statistics were performed.

For all analysis which examined before and after effects, tests for correlated samples were employed. On the other hand, when the experimental group was compared with the control group, tests for independent samples were employed.

Pre- and posttest differences for each corresponding pair of questions, as well as for the composite score, were compared for the attitude and knowledge methods. The scores were compared to determine whether the sum of all the knowledge or attitude responses changed. This procedure was not followed with the behavior data because it was not possible to compare scores with the statistic that was used for analyzing the behavior results.

Table 9. Statistics Used to Analyze Data for Study Objective One.

Statistic	Data Analyzed with Particular Statistic	Information Obtained
Frequencies	all	profile of data's distribution: pretest and posttest
T-tests for correlated samples	each knowledge question (test item) and knowledge score: control group pretest with posttest experimental group pretest with posttest	differences in means (correct) as a result of the treatment
T-tests for independent samples	each knowledge question (test item) knowledge scores, attitude scores: pretest control group with experimental group posttest control group with experimental group	differences in means (correct) between the two groups before and after the treatment
Wilcoxon Matched-Pairs Ranked Sign Tests	each ordinally measured attitude question (test item): experimental group's pretest with posttest control group's pretest responses with posttest responses	differences in medians (correct) as a result of the treatment

Table 9 (cont'd.)

Statistic	Data Analyzed with Particular Statistic	Information Obtained
Mann-Whitney U-tests	each ordinally measured attitude question (test item): pretest control group with experimental group posttest control group with experimental group	differences in medians (correct) between the two groups before and after the treatment
McNemar test	each nominally measured behavior question and attitude question: control group pretest with posttest experimental group pretest with posttest	differences in proportions as a result of the treatment

Knowledge

T-tests were calculated to compare differences of means from the pretest and posttest for the control group and the experimental group. To check the reliability of the correlated sample t-tests, t-tests for independent samples were used. They examined differences of means between the experimental group and the control group before the treatment and differences of means between the two groups after the treatment.

Attitudes

Three statistical tests were used to analyze the attitude data. On the ordinally measured data (questions 1 to 7 on the pretest and 1 to 8 on the posttest), Wilcoxon matched-pairs sign-ranked tests were used to compare the control and the experimental groups' and pre- and posttests. Mann-Whitney U-tests were employed to compare and contrast the experimental with the control group's ordinally measured attitude information. The attitudinal scores for the control group and experimental group were compared with each other using a t-test for independent samples. The third statistic, the McNemar test, was applied to the nominal level attitudinal data (pretest question 8 and posttest question 9). In this test, the control group's pretest and posttest results were compared to each other and the experimental group's data were treated in the same way.

Behavior

The McNemar test was used to examine changes before and after for the experimental and control groups for the corresponding question pairs (see page 51).

Program Content

The content checklists were reveiwed for major inconsistencies across programs, but no statistical tests were performed on this information.

Study Objective 2: Evaluation of Interpretive Methods

The multitrait multimethod matrix uses correlations to test if certain methods measure the same trait (Lenke and Wiersma, 1976; Campbell and Fiske, 1959; etc.). Pearson's correlations were calculated for the "scores" (see explanation in the beginning of this "Data Analysis: section) of the different measurement techniques (questionnaire, games, behavior intention story and observation). Pearson's correlations, rather than Chi square, Spearman's rank-order correlation coefficient or Kendall's tau were used, because the data were treated as interval level data. This treatment seemed justified as data that contain only two categories ("correct" and "incorrect" in this case) may be treated as interval level data (Nie, et al, 1975).

Scores from both the experimental and control groups were used to yield various correlations. For the experimental group, pretest and posttest scores were not correlated with each other because the treatment would affect the results. Instead, pretest scores were correlated with

pretest scores and posttest scores were correlated with posttest scores. Pretest scores were correlated with posttest scores for the control group. Table 10 shows the methods that were correlated with each other.

Table 10. Methods That Were Compared With Each Other for Study Objective Two.

Method A	Method B
Crossword Puzzle (Knowledge)	Questionnaire (Attitudes) Observation (Behavior) Behavior Intention (Behavior)
Matching Game (Knowledge)	Questionnaire Observation Behavior Intention
Questionnaire	Observation Behavior Intention

RESULTS AND DISCUSSION

This chapter of the document is divided into three major sections. The first section is an analysis of the changes in the methods and the reliability of the instruments. The second section describes the extent to which each program objective was met (achievement of external criteria), the changes before and after the program for each domain (knowledge, attitude and behavior), and the results from measuring respondents' level of schematic differentiation about dunes. The third major section describes the results from the correlations of the various methods.

Description of the Study Sample

The total study sample (n=112) consisted of 42% boys and 50% girls (8% did not respond). They ranged in age from eight to thirteen years; approximately 80% were ten to eleven years old. The sample participants came from three schools: Bennett School (Gary, IN), Edgewood School (Michigan City, IN), and Pieffer School (Schereville, IN). The experimental group consisted of 84 students from all three schools. Twenty-eight participants from Pieffer School made up the control group.

Analysis of the Methods and Instruments

Because the study design changed (see Methods) and the instruments were complex, an analysis of the methods and instruments is necessary to provide a suitable context from which to view the remaining results of this study.

Methods

Various changes in the methods resulted because the study was terminated before all data were collected. The total number of participants decreased from 484 to 112; this reduction effected all aspects of the study. To meet study objective one (evaluation of program effectiveness) both control group (n=24; original n=105) and experimental group (n=84; original n=379) were smaller. The number of individuals used to meet study objective two (evaluation of interpretive methods) was also decreased (experimental group n=21; original n=66), (control group n=6; original n=30). In addition, all experimental individuals' data were analyzed together even though individuals from different schools received slightly different treatments (Table 5). Limited staff also resulted in some changes in the methods. It was not possible to use observation as both a pretest and posttest measurement of behavior. The observation method was not checked for reliability.

Other factors effected the study design. Originally the study was designed so only first-time-visiting students would participate. The data show that over 33% of study participants indicated they had already visited INDU (Table 11). Hence, repeat visitation may have influenced participants' responses. It was also not possible to control the teachers' influence on participants.

The changes in the methods suggest that the results of this study must be viewed cautiously.

Table 11. Pretest and Posttest Attitude Responses for the Experimental and Control Groups Compared with External Criteria.

Objective Number(s): Content (Pretest Item #'s) (Posttest Item #'s)	Question Content	Stipulated External Criteria (% of total group)*	Experimental Group Pretest Observed (% of group)	Experimental Group Posttest Observed (% of group)	Control Group Pretest Observed (% of group)	Control Group Posttest Observed (% of group)
6: want to visit again (1,9,10,11, 12) (1,10,11,12, 13)	wanted to visit INDU again	80	67.9	97.4	50.0	91.7
	did not want to visit INDU again		3.6	2.4	8.3	0.0
	felt very good about INDU/nature parks		71.2	67.1	62.0	83.0
	felt good about INDU/nature parks		25.0	20.7	25.0	16.7
	felt neutral about INDU/nature parks		3.8	9.8	0.0	0.0
	wanted to return to to same nature park		76.2		65.2	
	did not want to return to same nature park		10.7		8.3	
	talk to families		84.5	83.3	66.7	95.8
	talk to friends		79.8	81.0	62.5	83.3
	visit with family/ parents		21.4	23.8	20.0	37.5
	visit with school		20.2	5.0	12.0	12.5
	visit with school, parents, friends		22.6	22.6	20.0	33.3
	visit with friends		8.3	10.0	4.2	0.0

*Appendix lists what respondents liked best and thought was the most important thing(s) they learned.
**External criteria based on program objectives, discussed on pages 33-34.

Table 11 (cont'd.)

Objective Number(s): Content (Pretest Item #'s) (Posttest Item #'s)	Question Content	Stipulated External Criteria (% of total group)*	Experimental Group Pretest Observed (% of group)	Experimental Group Posttest Observed (% of group)	Control Group Pretest Observed (% of group)	Control Group Posttest Observed (% of group)
7: feel important to protect plants/animals (5,6,7);(3,4,5)	felt very important to protect	80	67.5 60.5-75.3	59.6 48.7-70.0	56.8 50.0-62.5	47.9 39.1-54.5
	felt important to protect	0	21.4 13.6-32.1	20.6 17.5-21.2	19.4 12.5-25.0	17.5 9.1-26.1
8: enjoy the program (2,16,17)*	enjoyed very much	70	not applicable	60.0	not applicable	not applicable
	enjoyed			23.8		
	felt neutral			10.0		
9: want to prevent erosion (8); (9)	run on non-vegetated hill		47.4	51.9	29.0	43.5
	run on marked trails	50	26.9	35.4	33.3	34.5
	unsure		17.9	10.0	16.7	21.7
10: feel littering is bad (2,3,4);(6,7,8)	felt littering very bad		60.9 51.3-66.7	68.6 64.2-72.3	66.7 62.5-70.8	61.7 58.3-66.7
	felt littering bad	90	16.3 13.6-19.0	16.3 13.4-18.5	13.7 12.5-16.0	16.7 12.5-20.8

*Appendix lists what respondents liked best and thought was the most important thing(s) they learned.
**External criteria based on program objectives, discussed on pages 33-34.

Instruments

Reliability correlations were run with the control groups' data so the treatment did not effect the results of the correlations. Because no control group filled out a matching game knowledge instrument, it was not possible to calculate correlations for this instrument. Also, since the control group was small (see previous section) and no variance was found for their posttest behavior responses (pp. 94-96), a phi coefficient of association to check behavior pretest and posttest instruments' reliability could not be calculated.

Analysis of the crossword puzzle pretest and posttest items showed that the items were not strongly correlated (Table 12). In fact, for seven out of eleven questions it was not possible to calculate a correlation because the responses for either the pretest or posttest items did not vary. The largest positive correlation was for pretest-posttest item pair 5, 5 (.218). Three negative correlations also resulted (7 - 7, -.378; 9 - 10, -1.00; pretest score - posttest score, -.025). Because nearly the same wording and content were also used for the matching game it is probable that this instrument's test items would have shown similar correlation coefficients. Reliability correlations should range (approximately) from .7 to .9; hence, these results indicate that crossword knowledge instruments were not reliable.

As the instruments did not seem to measure the same information over time, the knowledge results must be viewed

Table 12. Reliability Correlations for Knowledge (Crossword Puzzle) Pretest and Posttest Items.

Pretest Question Number	Posttest Question Number	Pearson's Correlation Coefficient	N	P
1	8	not computable	2	
2	6	not computable	5	
3	4	not computable	11	
4	2D	not computable	6	
5	5	.218	10	.272
6	3	not computable	16	
7	7	-.378	8	.178
8	1	.178	13	.281
9	10	-1.00	2	.001
10	9	not computable	5	
11	2A	not computable	6	
score	score	-.025	21	.457

cautiously.

The reliability correlations between test items for the attitude instruments did not show strong correlations between pretest and posttest items and scores (Table 13). The highest correlation was found for pretest and posttest scores (.531) followed by pretest-posttest items pairs 5, 3 (.498) and 2, 8 (.448) and 7, 4 (.447). The lowest correlation was found for pretest-posttest item pairs 3, 7 (.004). One negative correlation was also found (item pair 1, 1; correlation -.269). However, reliability correlations should be higher than validity correlations (Kidder, 1981) and between .7 and .9. These data indicate that the attitude pretest and posttest instruments did not measure the same information over time. Therefore, the forthcoming results about participants' attitudes must be viewed with caution.

Study Objective 1: Evaluation of Program Effectiveness

Achievement of External Criteria

Knowledge: Pretest Responses

Approximately 71% of the experimental group were given the crossword puzzle and 29% did the matching game. All in the control group (100%) were given the crossword puzzle.

The conditions stipulated in the knowledge program objectives were not met before the program was carried out (Table 14), yet respondents did know some facts and concepts before the treatment. Over two-thirds (69.0%) of the experimental group who filled out the crossword puzzle knew that

Table 13. Reliability Correlations for Attitude Pretest and Posttest Items.

Pretest Question Number	Posttest Question Number	Spearman's Rank Order Correlation	N	P
1	1	-.269	22	.114
	2	.330	22	.067
2	8	.448	22	.019
3	7	.004	22	.493
4	6	.131	22	.282
5	3	.498	21	.011
6	5	.382	21	.044
7	4	.447	21	.022
score[1]	score[1]	.531	22	.006

[1]calculated with Pearson's Correlation.

plants hold the dunes in place to prevent erosion and that walking on plants can kill them (program objectives 1,2). Over half (57.5%) knew a common dune plant or animal (program objective 4). They knew least frequently (41.2%) that a dune was a hill of sand (Table 14). This result seems to indicate that the respondents knew that plants were needed to prevent erosion even though they did not actually know what a dune was. They were probably transferring previous knowledge about erosion on any soil type to erosion on the dunes. This interpretation must be made cautiously, given that t-tests (calculated with separate variances) showed a significant difference between the control and the experimental groups for items that asked about plants preventing erosion and people killing plants (pretest items) (Table 15). The result may suggest that the test items were not reliable or valid (rather than that the largest percent of the experimental group knew about plants preventing erosion and people killing plants) because the control group and experimental group responses should be similar for all pretest items, but in fact, were not.

The individuals who answered the matching game most frequently (34.8%) knew the name of a dune successional community (program objective 5) and least frequently (8.7%) what a dune was. Approximately one-fifth (21.7% and 19.6% respectively) knew that plants were important to prevent erosion, people could kill plants, and that plants and animals are protected in parks (program objectives 1, 2 and 3).

Table 14. Comparison of Experimental and Control Groups' Percent Correct Responses on the Crossword Puzzle and the Prestipulated Percent Correct.

Objective Number(s) Content (pretest item numbers); (posttest item numbers)	Stipulated External Condition % of Group Correct[1]	Experimental Group Pretest Observed % of Group Correct (mean)*	Experimental Group Posttest Observed % of Group Correct (mean)*	Control Group Pretest Observed % of Group Correct (mean)*	Control Group Posttest Observed % of Group Correct (mean)*
1,2: know plants prevent erosion; know walking on plants kills them (1,3,6,11); (9,5,4,2)	90	69.0	67.3	72.6	54.2
3: know plants/animals protected (5,9); (6,11)	80	48.1	48.3	42.8	33.3
4: know plants/animals have specific habitats (2,4,8); (7,3,1)	70	57.5	56.7	52.3	52.8
5: know a successional community (7); (5)	70	47.1	41.7	38.1	54.2
general: definition of a dune (10); (10)	not stipulated	41.2	61.7	33.3	37.5

*see Table 15 for the percent correct for each test item

[1] external criteria discussed on pages 33-34.

Table 15. Comparison of Experimental and Control Group Knowledge Pre- and Posttest Items.

Objective Content	Knowledge (Crossword) Test Item*	Pretest Experimental Group: % of Group Correct	Pretest Control Group: % of Group Correct	$t_{Observed}$**	(df)	P	Posttest Test Item*	Posttest Experimental Group: % of Group Correct	Posttest Control Group: % of Group Correct	$t_{Observed}$	(df)	P
know plants prevent erosion and walking on plants kills them	1	85.7	84.2	2.34	(46)	.024	9	55.8	12.5	2.63	(38)	.012
	3	71.4	72.5	1.61	(53)	.114	5	70.0	66.7	1.43	(43)	.160
	6	76.2	76.5	2.46	(44)	.018	4	83.3	79.2	-0.06	(35)	.955
	11	42.8	58.8	1.79	(32)	.083	2	60.0	58.3	-0.37	(30)	.712
know plants/ animals are protected	5	51.0	57.1	1.20	(36)	.238	6	53.3	45.8	-0.23	(19)	.817
	9	45.1	28.5	-0.43	(15)	.675	11	43.3	20.8	-0.45	(9)	.661
know plants/ animals have specific habitats	2	56.9	47.6	-0.08	(28)	.940	7	41.7	37.5	2.78	(31)	.009
	4	62.7	57.1	1.96	(35)	.058	3	65.0	47.1	1.00	(39)	.323
	8	52.9	52.3	-0.61	(17)	.547	1	63.3	79.2	1.32	(58)	.191
know a successional community	7	47.1	38.1	-0.44	(20)	.668	8	41.7	54.2	1.35	(41)	.184
know definition of a dune	10	41.2	33.3	-0.99	(13)	.349	10	61.7	37.5	-0.83	(12)	.424
Score				0.87	(37)	.388	Score			1.85	(51)	.069

*test items arranged according to program objectives (given on pages 33-34)
**t-values calculated using separate variance estimates (more conservative estimate)

This suggests some respondents knew specific information about park rules (protecting plants and animals) and appropriate behavior (not to walk on plants) even though they did not know what a dune was. The control group's responses were very similar to the experimental group's responses. Almost three-quarters (72.5%) of the control group knew that plants prevent erosion and walking on plants kills them. The fewest (33.3%) knew what a dune was (Table 16).

Group t-tests for independent samples showed significant differences between the control and experimental group for pretest items 1, 6 (Table 15). Walking on plants can kill them and that plants keep the dune from eroding are the content of pretest items one and six, respectively. These differences seem to indicate that the two groups of individuals were different in terms of these two questions, but were equivalent in terms of other (measured) knowledge.

Knowledge: Posttest Responses

The program objective conditions were also not achieved after the program. Similar to their responses for the pretest, experimental group respondents most often knew plants helped prevent erosion and that walking on plants could kill them which in turn would cause erosion. The second most frequent (61.7%) "correct" answer was knowing what a dune was. This was followed (56.7%) by knowing a common dune plant or animal's habitat and knowing that plants and animals were protected in the Lakeshore (48.3%). The percentage of "correct" answers about protecting plants

Table 16. Comparison of Experimental Group's Correct Responses on the Matching Game (Knowledge) with the Prestipulated Percent Correct

Objective Number(s): Content (pretest item numbers); (posttest item numbers)	Stipulated External Condition (% correct)[1]	Experimental Group+ Pretest Observed % of group correct (mean)*	Experimental Group+ Posttest Observed % of group correct (mean)*
1, 2: know plants prevent erosion; know walking on plants kills them (6,7,8,3); (8,3,9,4)	90	21.7	27.2
3: know plants/ animals protected (2,10); (6,1)	80	19.6	75.4
4: know plants/ animals have specific habitats (9,4,1); (10,7,3)	70	14.5	20.3
5: know a successional community (11); (11)	70	34.8	26.1
general: definition of a dune (5); (5)	not stipulated	8.7	17.4

+ - no matching tests given to the control groups
* - see Table for the percent correct for each test item
[1]external criteria discussed on page 33-34.

and animals, knowing a common dune plant's or animal's habitat, knowing about erosion, and knowing a successional community, stayed approximately the same before and after the treatment. This suggests that more respondents (41.2% before; 61.7% after) knew what a dune was after the program (Table 14).

Those who completed the matching most often "correctly" answered that animals and plants should be protected in the Lakeshore (75.4%). This was followed by knowing that plants hold the dune in place in order to prevent erosion and people can kill plants by walking on them (27.2%), and being able to identify a dune successional community (26.1%). These responses varied from those before the treatment. Before the treatment, respondents who filled out the matching game most frequently identified "correctly" a successional community but after the treatment the largest percentage answered that plants and animals should be protected. A greater percentage of respondents knew what a dune was after the treatment than before the treatment (8.4% and 17.4% respectively), although the fewest percentage of respondents knew this definition before and after the program. These data seem to suggest that respondents did not know specific definitions but were aware of rules and regulations (protection) and behavior that could cause damage (walking on plants will kill them, so erosion eventually occurs) after the program. The above information has important implications for resource managers. It suggests that ecological concepts and facts are not a necessary foundation for understanding appropriate park behavior (and the rules that govern park behavior). Therefore, if park management

objectives are geared to managing or changing behavior, programs should emphasize behaviors as they relate to resource management practices and use ecological concepts and facts to enrich the program rather than as the main focus. Also, since participants seemed to have a general understanding of appropriate behavior, but did not know what a dune was, participants need to learn that behavior and rules which they already know are applicable in the dune environment. Therefore, programs should focus on how these behaviors and rules apply to the dune settings. In addition, it might be possible for managers to assume that fourth-to-sixth grade students already know some appropriate behaviors.

The control group most frequently answered "correctly" that plants hold the dunes in place so erosion does not take place (54.2%) and the location of a particular successional community (54.2%). These responses were followed by knowing a common plant or animal's habitat (52.8%) (Table 13). These data differ somewhat from the experimental group's data as control group respondents indicated they knew more ecological facts than rules and behavior associated with rules. In fact, t-tests (using the more conservative separate variance estimates) showed significant differences between the experimental and control groups for posttest items 7 (name of a successional community) and 9 (definition of a dune). One would think that the content of the "correct" pretest responses would be similar for both groups but would

differ on the posttest. It is also expected that the control groups' answers would be similar on both the pretest and posttest. While these conditions generally seem to be met (Table 15) it is interesting to postulate why the discrepancy exists in the pretest. The small size of the control group (n=24) might account for some variation and hence not show representative responses. The differences may be a result of the science background of the control group respondents. For instance, the control group's teacher may have emphasized ecological facts in his/her class. (The control group's teacher also told them to stay on the trail and put their trash in the trash can (see "Behavior Posttest Responses" section.) The instruments may not be valid or reliable (Table 12) which implies additional research is necessary to develop reliable and valid instruments to use for evaluating interpretive effectiveness. However, the differences in posttest results between the two groups may be attributed to the effects of the treatment on the experimental group (see "Effects of Treatment - Knowledge Results" section).

Attitude: Pretest Responses

The external criteria were not met for pretest attitude responses, but interesting information is gleaned from examining these responses (Table 11).

Program Objective 6: Want to visit again. Almost 70 percent (67.9%) of the experimental group wanted to go (back) to INDU and 35.7% of them indicated they had already visited INDU. Half (50%) of the control group wanted to go

(back) to INDU but only 25% said they had visited before (Table 11). Since about twice as many respondents wanted to go back than had already visited (implying they had already been there) it is questionable that respondents really knew what, where, etc., INDU was. This may be particularly true for fourth to sixth graders who do not differentiate among parks (Indiana Dunes State Park is also nearby, for example). This result suggests that the instruments may not have validly measured whether respondents wanted to return to INDU. Hence, future investigators should spend additional time developing evaluation instruments that are known to be valid.

Respondents felt similarly about nature parks in general. Over three-quarters (76.2%) of the experimental group indicated that they liked going back to the same nature park more than once and 65.2% of the control group indicated likewise. In addition, both experimental and control groups (71.2% and 62.0% respectively) felt very good about nature parks in general and 25% felt good about nature parks in general. Most indicated they talked to their family and parents (84.5% experimental and 66.7% control group) after they visited a park. They also responded that they most often visited nature parks with their families and school groups (Table 11). These data seem to indicate that the respondents felt quite positive about both nature parks in general and INDU, even before they were exposed to the treatment.

Program Objective 7: Feel important to protect plants/animals. An average of 67.5% and 21.4% respectively of the experimental group felt it was very important and important to protect plants and animals. An average of 56.8% and 19.4% respectively of the control group felt it was very important and important to protect plants and animals (Table 11). Like respondents' views for objective one, nearly all those in the experimental and control groups had a positive attitude about protecting animals and plants before they received the treatment. Although the objective stipulated protecting plants and animals "in the Lakeshore," this phrase was not added to the instrument question for several reasons. The students would not necessarily know what the "Lakeshore" was, particularly in the pretest. If they had guessed what it was, the data would not have accurately indicated that they felt that protection was important in the "Lakeshore." Also, adding extra phrases to a lengthy instrument may have increased the non-response rate, particularly in the field. Because the phrase was deleted, the data were not a valid indication of participants feeling about protecting plants and animals in the "Lakeshore." Hence, future instruments should be developed which are more sensitive to subtleties such as feelings about certain objects "in the Lakeshore."

Program Objective 8: Enjoy the program. Not applicable on the pretest.

Program Objective 9: Want to prevent erosion. Almost half (47.4%) of the experimental group said they most wanted to run on a hill without plants and over one-quarter (26.9%) said they most wanted to run on marked trails. Almost one-fifth (17.9%) were not sure where they would most like to run. Seven and one-tenth (7.1%) percent indicated they would most like to run on a hill with plants.

The control group's responses were similar. One-third (33.3%) indicated they would most like to run on marked trails at a nature park, 29% said they would most like to run on a hill without plants and 16.7% were unsure. None of the control group individuals said they would most like to run down a hill with plants (Table 11).

This objective was difficult to measure because it is actually a behavior intention question instead of an attitude question. As such, the question attempted to measure what the respondents would intend to do. For example, if an individual wishes to help prevent erosion, he/she may still want to run on hills, particularly if running on hills is of higher priority to the individual than preventing erosion. Before the program, only 7% of the experimental group and none of the control group said they wanted to run down a hill with plants, yet running on a hill with plants may be acceptable, especially if the respondents perceived plants as "grass." But even before the program, these data indicate that most respondents wanted to run down hills without plants or on trails. These behaviors would not

cause erosion. It seems then that respondents would want to help prevent erosion. This interpretation is reinforced in the knowledge results: Respondents knew that walking on plants could eventually cause erosion, i.e., they had general knowledge about proper behavior.

Program Objective 10: Feel littering is bad. An average of 60.9% and 10.8% respectively of the experimental group felt that littering was very bad and bad (Table 11). The control group felt similarly about litter: An average of two-thirds (66.7%) felt that littering was very bad and 13.7% felt littering was bad.

As with attitude program objectives six and seven, the respondents were strongly oriented toward the desired attitude - in this case, negative feelings about littering.

Attitude: Posttest Responses

The stipulated program objective percentages were not attained for the posttest attitude questionnaire items. Perhaps the most important program objectives from a park manager's perspective, numbers six and eight, were closest to the desired condition. About 50% of responses were consistent with the program objectives for the other three objectives.

Program Objective 6: Want to visit again. Program objective six (that participants want to return to INDU) was met for two of the five questions that measured the objective (Table 11). Almost all (97.4%) of the experimental group and 91.7% of the control group wanted to return to

INDU. Over two-thirds (67.1%) of the experimental group said they felt very good about INDU before they came and one-fifth (20.7%) said they felt good. Eighty-three percent of the control group said they felt very good about INDU before they came and 16.7% said they felt good. Just under one-tenth (9.8%) of the experimental respondents and none of the control group said they felt neutral about INDU before they came.

In addition, over 50% of respondents in both the experimental and control group said they would come with their family or school if they returned. One-third (33%) of the control group and 22.6% of the experimental group said they would bring friends, family and school class with them if they returned to INDU. Most of the experimental group said they would talk to their family (83.3%) and their friends (81.0%) about their field trip. Most of the control group also indicated they would talk to their families (95.8%) and friends (83.3%) about their trip to INDU.

Mann-Whitney U- tests indicated no statistical differences between the experimental and control groups' post-test attitudes about returning to INDU (Table 17) indicating both groups felt similarly about returning to INDU. However, one would expect the experimental group to be more likely to want to return than the control group, as the experimental group was exposed to the treatment. Only small differences existed between the two groups and in some cases the control group answered more favorably to the

Table 17. Comparison of Control Group and Experimental Group Pretest and Posttest Attitude Items.

Pretest					Posttest				
Attitude Test Item*	Experimental Group: % of Total Group Correct+	Control Group: % of Total Group Correct+	Z Observed (n)**	P	Attitude Test Item*	Experimental Group: % of Total Group Correct+	Control Group: % of Total Group Correct+	Z Observed (n)**	P
1	96.2	87.0	-0.660(58)	.509	1	87.8	99.7	.950(60)	.342
					2	83.3	100.0	.950(58)	.365
5	66.7	50.0	1.559(58)	.119	3	70.0	50.0	1.65 (59)	.199
6	75.3	62.5	1.080(58)	.280	5	60.0	54.5	1.36 (58)	.173
7	60.5	58.0	-0.217(58)	.828	4	48.7	39.1	.435(57)	.665
2	80.1	75.0	-0.969(57)	.333	8	86.6	66.7	.707(60)	.479
3	83.6	82.7	-0.852(56)	.394	7	85.4	58.3	1.198(60)	.231
4	82.7	83.3	-1.376(58)	.169	6	82.7	58.3	.405(60)	.685
Score			0.021(59)	.983	Score			1.115(60)	.265

*items arranged according to program objectives on page 34, program objective 8 not applicable; program objective 9 not analyzed in this test

+percent correct based on objectives

**Z scores calculated for Mann-Whitney U-tests when n>30.

questions. More of the experimental group (97.4%) said they wanted to visit again than the control group (91.7%), although more of the control group said they felt very good about INDU (83%), indicated they would talk to their families (95.8%) and friends (83.3%) than the experimental group. These data suggest that an interpretive program is not necessary for respondents to feel good about a certain park (INDU) and talk to friends and family about their visit to INDU. This has important management implications. First, if people feel good about visiting a park, other factors besides interpretive programs may give them those feelings. These factors include what they hear about the park from other people, including school, family and friends, previous experiences at the specific park or other similar parks, the management practices which park personnel employ and the meaning of the park that an individual construes based on his/her background, cultural upbringing, etc. (i.e., the social meaning). Second, control (and experimental) respondents indicated they would talk to their families and friends about the park; these conversations may powerfully persuade others to come to the park or not. If respondents had a bad time, and talked to others about it, chances are the other people would not visit the park either. Therefore, it is important that visitors view all aspects of park operation (rules, regulations, resource management practices, as well as interpretive programs) favorably. For instance, management practices for appropriate park behavior should

meet users' needs while simultaneously protecting the resource. This suggestion implies that park personnel must "know" their users/visitors in order to cater to their needs.

The differences between the two groups and the meanings given for these differences must be cautiously interpreted as various factors may have influenced the difference. The two groups' responses may have been similar because the control group received a "partial treatment" even though they were not exposed to the interpretive program. The "partial treatment" included the bus trip through the country, the drive through the park, time away from school, etc. Hence their responses about INDU and returning to INDU may have been influenced by this "partial treatment." This suggests that in future studies the control groups' posttest should be administered at the schools. The differences between the two groups may be due to chance since, Mann-Whitney U-tests indicated no statistically significant differences between the groups (Table 17). The students in either group may not have had any feelings about INDU but the instruments did not measure this possibility.

Program Objective 7: Feel important to protect plants/animals. The stipulated 80% criteria for protecting plants and animals was not achieved. An average of 59.6% of the experimental and 47.9% of the control group felt that it was very important to protect plants and animals. In both groups, nearly one-fifth (20.6% experimental and 17.5%

control) felt that it was important to protect plants and animals (Table 11).

For both attitude strengths (important and very important) fewer control group respondents answered "correctly" than the experimental group (yet not statistically significantly fewer) (Table 17). These data seem to suggest that the program did effect the experimental group's attitudes about protecting plants and animals. Knowing that plants and animals are protected in the Lakeshore was one of the largest changes found in respondents' (who completed the matching) knowledge (see "Knowledge Measurements: Posttest Responses" section). Hence, it seems logical and consistent that the program also effected the attitudes that respondents had about protecting plants and animals.

Program Objective 8: Enjoy the program. Program objective 8 was nearly attained for the experimental group (not applicable for the control group). Sixty percent of the experimental group said they enjoyed the program very much (stipulated 70%) and 23.8% said they enjoyed the program. Respondents also indicated what they thought was the most important thing they learned. Their comments ranged from "everything" to "don't walk on the plants" to "how some plants look" to "don't destroy nature." They said what they liked best about the program. These comments ranged from running and walking on the hills and dunes themselves to water and the lake. Appendix E lists what respondents thought was the most important thing(s) they learned and

what they liked best about the program.

These comments suggest that the program was successful even though all the knowledge and attitude criteria were not satisfied. In fact, none of the experimental respondents failed to answer what they thought was the most important thing they learned (question 16) and what they liked best (question 17). Respondents most frequently answered "how some plants look" (10) as the most important thing they learned followed by "plants and animals" (8), and then "don't walk on the grass" (6) and "animals" (5). By far the aspect of the trip they liked best was "walking and running on dunes" (24) followed by dunes (8). Control respondents most often answered "climbing, running, etc., on the dunes" for posttest item 10A ("what would you most like to do?"). Also, before individuals received the treatment, the majority (both control and experimental) indicated they would most like to run on either a hill without plants or a marked trail (see "Attitude Pretest Responses: Program Objective 9"). This information, that they liked running down the hills, is important for future management practices. It suggests that a dune should be designated for running and climbing on. If one dune is set aside for this purpose, people could run and climb on this dune and not disturb other vegetated dunes and/or dunes that are blownout or eroding. By adopting this management practice, managers could protect the resource and meet visitor needs.

Program Objective 10: Feel littering is bad. Over two-thirds (68.8% and 61.6% of the experimental and control groups, respectively) said they felt littering was very bad. An average of 16.7% and 16.3% of the control and experimental groups respectively indicated they felt littering was bad (Table 11). However, a small difference (not statistically significant; Table 17) was found between the control and experimental groups' feeling that littering was very bad, suggesting that the experimental group had stronger feelings after the program. This trend is also found in the experimental group when the difference before (60.4% felt littering was very bad) and after (68.6% felt littering was very bad) treatment are examined (Table 11).

Behavior: Pretest Responses

Program Objective 11: Stay on trail. Almost all of the experimental group (96.2%) said they would put a candy wrapper from lunch into a trash can (question one) and 92.5% said they would put trash from a snack into a trash can (question four). The control group responded similarly. Nearly all (95.2%) said they would put their candy wrappers in a trash can. Ninety percent said they would put litter from a snack in a trash barrel (Table 18). These data suggest that respondents knew proper litter-disposal behavior.

Program Objective 12: Stay on trail. This objective was not measured because insufficient numbers of staff were available to observe.

Table 18. Control and Experimental Groups Pre- and Posttest Behavior Responses Compared With External Criteria.

Objective #; Content (pretest item #'s) (posttest item #'s)	Pretest Item Content / Posttest Item Content	Stipulated External Criteria	Experimental Group+ Pretest Responses (% of group)	Experimental Group+ Posttest Responses (% of group)	Control Group** Pretest Responses (% of group)	Control Group** Posttest Responses (% of group)
11: stay on trail (2,3);	pre: stay on trail	80	61.1		65.0	
	follow trail		15.6		37.5[1]	
	ask ranger (Q3)		40.5		45.0	
	post: main trail			28.6		100.0
	unmarked trail			47.0		
	off all trails			23.6		
12: stay on trail*		20				
13: properly dispose of trash (1,4);	pre: trash can (Q1)(Q4)	50	94.4[2]		92.6[2]	
	post: candy and wrapper in pocket			66.7		
	wrapper in pocket			20.0		
	wrapper on trail			13.0		
	trash can					100.0

*not measured due to insufficient staff
+n=16
**n=6
[1]this figure is the average percent of two possible choices
[2]this figure is the average percent from two questions

Program Objective 13: Properly dispose trash. The prestipulated external conditions were achieved for this objective. Three-quarters (75.3%) of the experimental group said they would "stay on the trail and remain quiet" if there was an interesting animal (question two). An additional 22.2% indicated they would take a trail to the animal. Only one respondent said he/she would run straight to the animal (and off the trail). Almost half (46.8%) of the experimental group also indicated they would view a plant from where they were and again only one respondent said he/she would run off the trail to the plant. Just under half (40.5%) said they would ask the ranger if they would see other similar plants and 8.9% said they would follow a trail to see the plant from the other side (question three) (Table 18).

The control group's responses were like the experimental group's. For pretest question two, 65% said they would stay on the trail and another 30% said they would take a trail to the animal. Almost half (45%) answered pretest question three "look at the plant from where you are" and an additional 45% said they would ask the ranger if they would see other similar plants later on. Ten percent (n=2) said they would pick the plant (Table 18).

The responses to littering and staying on the trail behaviors indicate the program objectives were met before the program and that the majority (for both the control and experimental groups) knew what appropriate behavior was.

This finding is reinforced by the attitude and knowledge responses that were given (see previous sections).

Behavior: Posttest Responses

Eighty-one percent of the experimental group (n=68) were not observed and twenty percent (n=16) were observed (see "Study Design" and "Study Sample" sections in Methods chapter).

Program Objective 11: Stay on trail. The stipulated external condition (80% stay on trail) was not satisfied for this objective on the posttest (Table 18). Almost one-third (28.6%) went on the main trail, nearly half (47.6%) went on an unmarked trail to meet the interpreter and 23.8% did not travel on any trail. Six individuals were observed in the control group and all six went on an unmarked trail to meet the interpreter. The teacher who was with the control group told the students to stay on the trail. But, participants still went on an unmarked trail. This result suggests that participants (and the teacher) did not differentiate between marked and unmarked trails (i.e., they were not aware the trails were trails which park personnel established). This suggests that park personnel need to eliminate unmarked trails so visitors do not need to differentiate between marked and unmarked ones, and clearly mark the trails on which they want people to walk (or run).

The unmarked trail was clearly the most direct route from point A to B (i.e., the shortest distance between two points). All the control group and the largest percentage

of the experimental group followed this path even though this was not the desired behavior. These data seem to indicate that respondents generally stayed on a trail, especially when the trail was a direct route to the desired destination. These results have important management implications. They suggest that most visitors will stay on established trails, particularly when the established trails are a direct route to where people want to go. Therefore, park personnel need to observe travel patterns of visitors in various areas of West Beach (and the park in general) to determine the existing direct routes and design main trails based on visitor travel patterns. Other unmarked trails should then be replanted (and fenced off as necessary to allow regrowth).

Program Objective 13: Properly dispose trash. Fifty percent (stipulated external condition) of participants did not correctly dispose their trash (Table 18). Nearly two-thirds (66.0%) of the experimental group put their wrapped candy into their pockets. Twenty percent put only the wrapper in their pocket and 19% dropped the wrapper on the trail. All (100%) of the control group who were observed put their trash in the trash can. The control group probably behaved this way for two reasons. First, their teacher told them to put the wrapper in the trash can. Second, the respondents were given their candy in plain view of a trash can. (However, this was done because it was the best place for the researcher to administer the control group's posttest.

Also, it was the best place to position the group so they were not standing right next to a trail that they might use.)

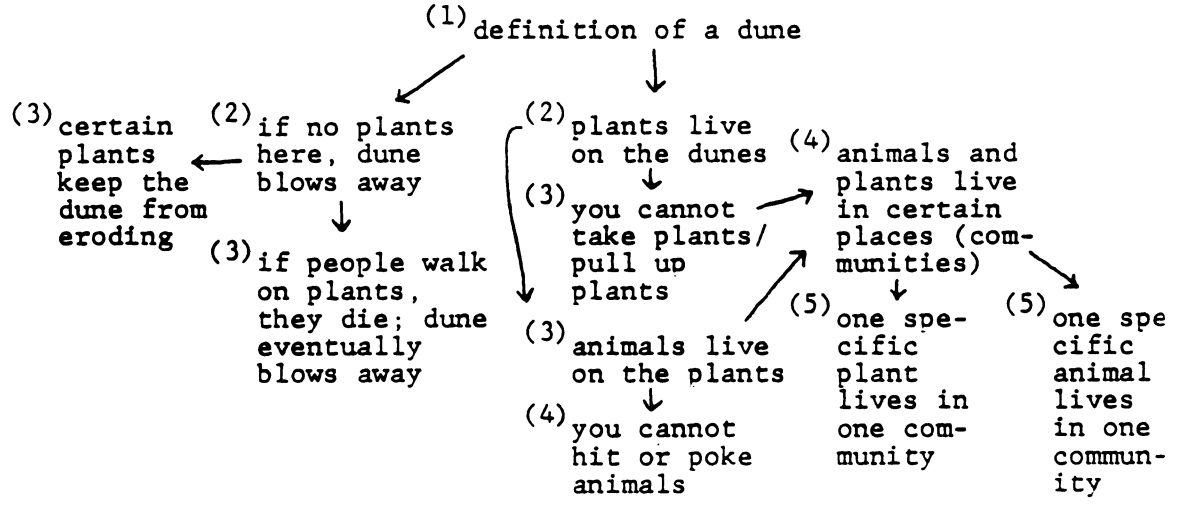
In other words, participants knew the correct behavior (and attained the stipulated criteria on the pretest) but did not necessarily behave correctly.

Measurement of (Knowledge) Schematic Differentiation

The knowledge questions were designed to measure participants' schematic differentiation about dunes. Figure 10 shows the relationships among the questions and how they are more or less complex relative to the other questions. Table 19 details the question content, its level of differentiation and the pre- and posttest question numbers that correspond to each level of differentiation for both the matching game and crossword puzzle.

Pretest Responses

Respondents percentage "correct" varied with the questions' level of differentiation for both the control and experimental groups (Figure 11, Table 20). For instance, experimental individuals who completed the matching responded "correctly" to the greatest percentages (37%) of answers at level four, followed by level one (34.8%), then three (24.6%), then two (17.4%), and finally level five (10.9%). The crossword puzzle respondents answered a greater percentage "correct" at level five (57.8% experimental group, 55.1% control group) than at level four. Also, the percentage "correct" at level one for the crossword puzzle respondents was less than both levels two and three.



Key - Arrows indicate the flow from less to more differentiated. In other words, it is necessary to know the information at the blunt end of the arrow before one can logically know the facts or concepts at the arrow's pointed end.

The numbers indicate the level of differentiation. For example, a low number represents a low level of schematic organization and a high number, a high level.

Figure 10. Relationships Among Knowledge Questions to Show Levels of Schematic Differentiation.

Table 19. Content of Knowledge Instruments for Each Level of Schematic Differentiation and Instrument Questions that Correspond to the Specific Levels.

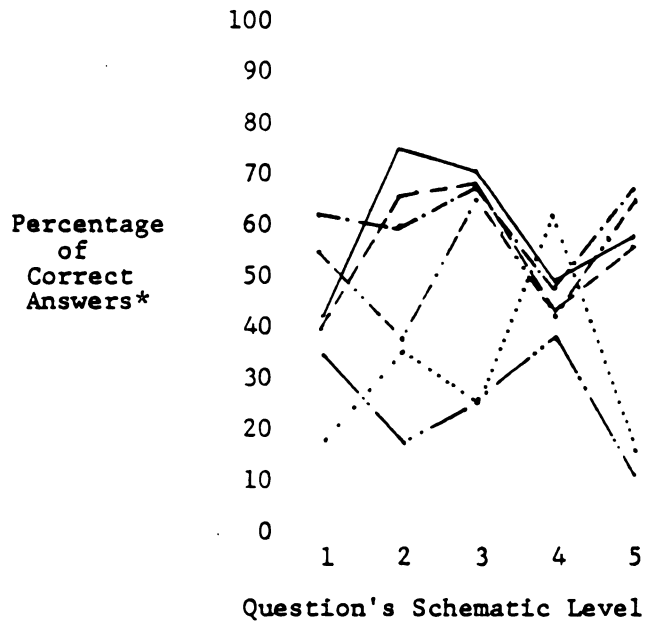
Question Content*	Corresponding Question Numbers on Instruments			
	Crossword		Matching	
	Pretest	Posttest	Pretest	Posttest
(1) question about what is a dune	10	9	5	5
(2) if no plants grow here, dunes blow away	1	8	6	8
(2) plants live on the dunes	11	2 down	3	9
(3) if people walk on plants, they die	6	3	7	2
(3) animals live on the dunes	2	6	9	10
(3) question about specific plant that keeps dunes from eroding	3	4	8	4
(3) question about protecting and disturbing plants	9	10	2	6
(4) question about protecting and disturbing animals	5	5	10	1
(4) question about identifying a successional community	7	7	11	11
(5) question about a specific animal's habitat	8	1	4	7
(5) question about a specific plant's habitat	4	2 across	1	3

*question content listed from top to bottom are less to more differentiated

Table 20. Percentage of Correct Responses for Each Level of Schematic Differentiation for the Pretest and Posttest.

Schematic Level of Differentiation	Test Item Number (pretest); (posttest)	Experimental Group Pretest: Mean % Correct	Experimental Group Posttest: Mean % Correct	Control Group* Pretest: Mean % Correct	Control Group* Posttest: Mean % Correct
			Crossword Puzzle		
1	(10);(5)	41.2	61.7	38.1	54.2
2	(1,11);(8,9)	70.6	57.5	64.3	35.4
3	(2,3,6);(2,4,10)	68.6	65.0	65.1	61.1
4	(5,9,7);(1,6,11)	47.7	46.1	41.2	40.3
5	(4,8);(7,3)	57.8	64.2	55.1	60.5
			Matching Game		
1	(5);(10)	34.8	17.9		
2	(6,3);(9,2)	17.4	32.6		
3	(7,8,9);(7,5,4)	24.6	24.6		
4	(2,10,11);(6,11,8)	37.0	58.0		
5	(4,1);(3,1)	10.9	15.2		

*No control group filled out the matching.



*these points were calculated by taking the mean of the percentages that existed for a particular differentiation level.

- - - pretest crossword - control
——— pretest crossword - experimental
—·—·— pretest matching - experimental
—·—·— posttest crossword - control
—·— posttest crossword - experimental
····· posttest matching - experimental

Figure 11. Percent of Correct Responses for Levels of Schematic Differentiation for the Control and Experimental Groups' Pretest and Posttests.

It would be expected that as the questions' level of schematic differentiation increases, the respondents' percentages of correct answers would decrease (i.e., an inverse relationship should exist between question's level of differentiation and respondents' percent of correct answers). This relationship should exist because the more complex answers would be more difficult for the majority to answer "correctly" (Figure 12). These data suggest that an inverse relationship did not exist. Rather, these data suggest that often times respondents knew certain (often behavioral) facts but did not know the underlying concepts that the facts were based upon. This trend is also evident when one examines the individuals responses by knowledge program objective (see "Knowledge: Pretest Responses" section).

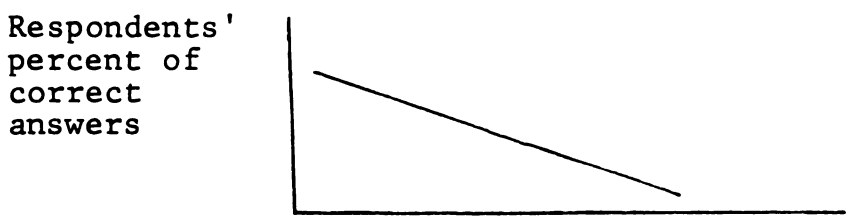


Figure 12. Expected Pretest Relationship Between Percentage of Correct Responses and Schematic Level of Differentiation About Dunes.

Posttest Responses

Neither the control nor experimental groups' responses followed any pattern of schematic differentiation (Table 20, Figure 11). More of the experimental individuals, who did the matching game, responded "correctly" at level two (32.6%)

than level one (17.9%) and fewer responded "correctly" than at level three (24.6%) and four (58.0%). At level five, fewer responded "correctly" than at level four (58%). The experimental individuals who did the crossword puzzle responded "correctly" most often for level three (65.0%), followed closely by level five (64.2%) and level one (61.7%) and then level two (57.5%). Those in the control group answered "correctly" for level three more often (61.1%) than any other level. This was followed by level five (60.5%) and then level one (54.2%).

An inverse relationship between percentage of correct answers and the questions' level of differentiation would be expected (see previous section and Figure 12). After the treatment two types of inverse relationships could exist. First, the slope could remain unchanged from the pretest slope (case A) and a greater percentage of "correct" answers would be found at all differentiation levels (i.e., the posttest line that showed percent correct would be parallel to the pretest line that showed percent correct). Second, the line could move toward the horizontal as more respondents answered the more differentiated questions "correctly" (case B) (Figure 13). However, the data indicated no relationship.

These results may be explained in several ways. When the crossword puzzle was employed, respondents (both control and experimental) answered correctly most often for level three (i.e., that plants keep the dune from eroding, that stepping on plants can kill them, that animals live on

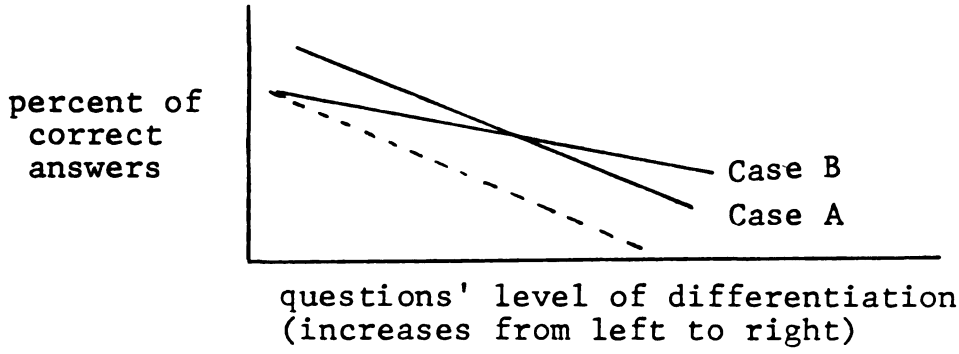


Figure 13. Expected Posttest Relationship Between Percentage of "Correct" Answers and Schematic Level of Differentiation. (- - - Indicates Pretest Relationship.)

plants, and that people cannot take or pull up plants. This level may have had the highest precent correct because four different concepts were assigned to level three and two or fewer made up the other levels, so respondents had the greatest opportunity to answer these questions correctly. This suggests future study instruments should have equal numbers of questions/concepts for each level. Those who completed the matching most frequently answered level number four "correctly." The differences in correct answers from the two instrument types was probably due to the nature of the levels of differentiation or the nature of the instruments (see "Discussion" section). Although certain concepts were assigned to a given level, the concepts at each level are not necessarily related. For instance, just because an individual answered "correctly" that certain plants keep the dunes from eroding (level 3), it does not necessarily follow that he/she would know that animals live on plants (level 3). In addition,

because the levels were developed around the program objectives, certain behaviors that respondents were to know were included. Individuals may have known certain behaviors without knowing specific ecological facts (e.g., they can transfer knowledge about behaviors to the dune environment). This explanation suggests that future researchers should design instruments which measure only levels of schematic differentiation of ecological facts and concepts and not levels of schematic differentiation which include behaviors. The way that the levels of differentiation were designed may be incorrect. For example, a definition of a dune was the simplest level of differentiation while knowing something about suitable behavior on the dunes was a more complex differentiation level. Appropriate behaviors, however, (staying off plants, not hurting animals, etc.) are simple cultural knowledge.

This explanation further suggests that levels of differentiation should be limited to only ecological information.

Treatment Effects

Since the program objectives were not satisfied (except behavior: pretest responses, which are the opposite of the expected), this section is a description of the statistical changes that resulted from the effect of the program (i.e., treatment).

Treatment Effects on Knowledge

For the most part, changes in participants' knowledge was not statistically significant as a result of exposure

to the program. None of the responses on the crossword puzzle changed significantly. Four significant changes were found for the matching game responses (pairs 2,6; 3,9; 10,1 and the score) (Table 21). Two of the three matching question pairs that changed had to do with plants (2,6;10,1). Question pair two and six asked from where people were allowed to take plants (protecting plants). Question pair three and nine asked where plants live. Question pair ten and one dealt with protecting animals. These changes suggest that as a result of the program (experimental) participants learned that protecting plants and animals was important. Also, group t-tests (calculated with separate variance estimates), showed a significant difference between the control and experimental group on posttest item 9 (Table 15). However, one must interpret this cautiously as reliability checks for all three question pairs indicated that pretest and posttest questions were not strongly correlated so the question pair responses may have changed because they do not measure the same thing over time (Table 12). These results are further complicated by using paired sample (or correlated) t-tests. Paired-sample t-test analysis assumes that the individuals are matched on some criteria or the same individual is tested over time. This analysis also assumes that the correlations between the two items will be large (in magnitude) and positive (Nie, et al, 1975; Fergeson, 1971). Although the same individuals were tested over time, the correlations were not expected. This again

Table 21. Treatment Effects on Knowledge: Pretest/Posttest Changes for the Experimental and Control Group.

Question pair** (pretest, posttest)	Question Content	Experimental Group Pretest: Mean % Correct (for group)+	Experimental Group Posttest: Mean % Correct (for group)	df	$t_{Observed}$	P	Control Group* Pretest: Mean % Correct (for group)	Control Group* Posttest: Mean % Correct (for group)	df	$t_{Observed}$	P
				Crossword Puzzle							
1,8	dunes without plants blow away	85.7	55.8	29	0.00	1.00	84.2	12.5	1	99.00@	99.00
2,6	animals live on dunes	56.9	41.7	22	1.28	.213	47.6	37.5	4	2.45	.070
3,4	specific plant keeps dune in place	71.4	70.0	33	0.81	.422	72.5	66.7	10	99.00	99.00
4,2A	plant in a successional community	62.7	60.0	25	1.44	.161	57.1	58.3	5	99.00	99.00
5,5	animals are protected	51.0	53.3	31	1.00	.325	57.1	45.8	9	0.56	.591
6,3	walking on plants kills them	76.2	83.3	42	0.70	.486	76.5	79.2	15	99.00	99.00
7,7	name of successional community	47.1	41.7	20	0.33	.748	38.1	54.2	7	1.43	.197
8,1	animal in a successional community	52.9	63.3	28	-1.28	.212	52.3	79.2	12	0.00	1.00
9,10	plants are protected	45.1	43.3	23	1.45	.162	28.5	20.8	1	0.33	.795
10,9	wind blows sand into dunes	41.2	61.7	20	1.00	.329	33.3	37.5	4	1.50	.208
11,2D	plants live on dunes	42.8	65.0	22	-1.45	.162	58.8	47.1	5	-1.00	.363
Score				50	-.04	.968			20	0.37	.718
				Matching Game							
1,3	plant in a successional community	0.0	17.4	15	1.81	.091					
2,6	plants are protected	17.4	65.2	21	4.42	.000#					
3,9	plants live on dunes	21.7	47.8	20	2.10	.049#					
4,7	animal in a successional community	21.7	13.0	17	-0.89	.383					
5,5	dunes are sand hills	8.7	17.4	20	0.54	.596					
6,8	dunes without plants blow away	13.0	17.4	19	1.62	.121					
7,2	walking on plants kills them	4.3	21.7	21	1.58	.129					
8,4	specific plant holds dune in place	47.8	21.7	20	-1.52	.145					
9,10	animals live on dunes	21.7	30.4	20	0.32	.752					
10,1	animals are protected	21.7	82.6	20	5.65	.000#					
11,11	name of successional community	34.8	26.1	18	-1.23	.236					
Score				21	-3.59	.002#					

*A control did not fill out the matching game.
+percent correct, as determined by program objectives (pages 33-34)
**parallel test items on the pre- and posttest instruments
#statistically significant
@incomputable

suggests the results obtained from the correlated t-tests are somewhat questionable. For example, some seemingly large differences were not significant and some small differences were significant (Table 21).

As few statistically significant changes in knowledge were found, it is necessary to investigate the more subtle changes that may have resulted from exposure to the program. Although administrators often think that results of program evaluation will show large effects, this case rarely prevails, and hence the smaller effects of the program must be examined (Rossi, 1972; Anderson and Ball, 1978).

The experimental group who filled out the crossword puzzle, answered "correctly" on the posttest a greater percentage of the time than the control group for all program objectives except five, identifying a particular dune successional community (Table 15). This seems to suggest that the program did have some impact on the experimental individuals. The largest difference was found on the question that asked what a dune was. The second largest difference between the pretest and posttest matching was also on the question that asked what a dune was (experimental group only). In addition, the responses from the pretest and posttest matching showed the largest difference for knowing about protecting plants and animals in the Lakeshore (program objective 3) (19.6% "correct" on pretest, 75.4% "correct" on posttest). This information reveals that defining a dune and protecting plants and

animals seems to have been the most important knowledge information (that was measured) that the program provided. (However, this interpretation must be cautiously applied as not all the program information was measured and the instruments did not seen to be reliable.)

Treatment Effects on Attitudes

Statistics for independent samples indicated that the two groups' attitudes did not differ significantly either before or after the treatment (Table 14). Therefore, more detailed inspection of changes may provide information about the program's effect on attitudes.

The Wilcoxon match pairs signed-rank test indicated four significant changes in the experimental group and one significant change in the control group (Table 22). Three question pairs that dealt with protecting plants and animals changed significantly (question pairs 5 and 6 [for both experimental and control group] and question pair 7 and 4 [for the experimental group] not disturbing plants and animals). In addition, previous analysis (see "Treatment Effects on Knowledge" section) showed that significantly more respondents knew that plants and animals were protected after they had the treatment than before. The results from both knowledge and attitudes measurements appear to suggest that more participants felt protecting plants and animals was important after being exposed to the program. Yet, the shift in question pair seven and four was in the opposite of the desired direction which suggests that respondents'

Table 22. Treatment Effects on Attitudes: Pretest/Posttest Changes for the Experimental and Control Groups.

Question pair (pretest posttest)	Question Content	Experimental Group Pretest: % correct (for group)	Experimental Group Posttest: % correct (for group)	N	$z_{Observed}$	P	Control Group Pretest: % correct (for group)	Control Group Posttest: % correct (for group)	N	$z_{Observed}$	P
1,1	feelings about nature parks	96.2	87.8	79	-1.46	.144	87.0	99.7	22	-1.28	.203
1,3	feelings about program	96.2	83.3	77	-2.39	.017	87.0	100.0	22	-.314	.753
2,8	feelings about litter	80.1	86.6	79	-2.16	.030	75.0	66.7	22	-.770	.442
3,7	feelings about litter	83.6	85.4	78	-.706	.480	82.7	58.3	22	-1.16	.248
4,6	feelings about litter	82.7	82.7	79	-.569	.569	83.3	58.3	22	-1.07	.285
5,3	feelings about protecting plants/ animals	66.7	70.0	78	-.821	.412	50.0	50.0	21	-1.05	.294
6,5	feelings about protecting plants/ animals	75.3	60.0	78	-2.96	.003	62.5	54.5	21	-2.27	.023
7,4	feelings about protecting plants/ animals	60.5	48.7	76	-3.50	.000	58.0	39.1	21	-1.73	.083
8,9	feelings about running on marked trails	26.9	35.4	pre-21 post-28	N.A.	.055	33.0	34.8	8	N.A.	1.00

attitudes changed in the negative direction. Given the above results and the reliability of the attitude instruments (Table 13), it is difficult to interpret the shift in participants' attitudes about protecting plants and animals. Therefore, the conclusion (that protecting plants and animals was the most prevelant positive attitude shift), must be viewed carefully.

Question pair one and two (feelings about the program) changed significantly for the experimental group. This result has important management implications. Respondents felt statistically significantly more negative about the program after the program than before it. In addition, somewhat fewer experimental respondents felt very good about nature parks (program objective 6) after the treatment than before. If people feel worse about nature parks AFTER they have participated in a program, they will be less likely to return to the park and also less likely to take part in future interpretive activities. Since repeat visitation is a desirable action for many reasons (e.g., helps build a strong support base for the park and helps management get to know the visitors better, etc.) personnel must try to make certain that people feel good about parks (and INDU in particular). Some aspect of the program or field trip in some way or another affected how respondents felt about parks, and unfortunately, affected them in a negative way. This suggests that future programs should try to instill positive feelings about parks to the participants and that

future research should aim to find out why these negative feelings may have developed.

One question pair (two and eight) out of three was statistically different for the experimental groups feelings about littering. In addition, the experimental group had stronger antilittering attitudes than the control group on the posttest. These data seem to suggest that experimental participants' had stronger attitudes about littering because they were exposed to the program.

For program objective seven (feel important to protect plants/animals), nine (want to prevent erosion), and most of the questions that measured program objective six (want to visit again), the control group answered consistently less often with the objectives than the experimental group did. This suggests that the program did effect (but not significantly) the experimental group. Their attitudes about protecting animals and plants were stronger than the control group's. The experimental group also said more often than the control group that they would most like to run on a hill without plants and they said this more frequently on the posttest than on the pretest. However, for program objective seven, the experimental groups said more frequently on the pretest that protecting plants and animals was important and very important than on the posttest. This difference is the opposite of what was expected. It may have been due to chance (because the difference was not significant). It may also be due to the instruments reliability (see "Analysis

of Methods and Instruments" section) or it may be because the program affected participants undesirably. Hence, this result must be carefully interpreted.

Treatment Effects on Behavior

The experimental participants' littering behavior (program objective 13) changed significantly before and after the program. This change was not in the desired direction. Before the treatment the majority indicated they would dispose of their trash in a trash can (Table 18). In reality, many participants did not actually behave this way. This result may seem odd given that the majority in both the control and experimental groups said they felt littering was bad or very bad. Heberlein (1971), however, found no relationship between antilittering attitudes and actual behavior. Other research (Clark, Hendee and Burgess, 1972; Clark, et al, 1971; Powers, et al, 1973) found the best way to encourage antilittering behavior was by offering a small incentive or reward for the behavior and this study did not offer any incentive.

One question about trail behavior (program objective 11) differed significantly but the other did not (Table 23). In both groups, question two was statistically significantly different from trail observation, but question three was not. The results for question three can be explained from the statistic that was used to calculate differences before and after the treatment, the McNemar test. This test requires that the variables be dichotomized. In other words, one

Table 23. Changes in Behavior for the Control Group and Experimental Group[1].

Behavior Intention Question Number	Behavior Observed	Experimental Group X^2	N	P	Control Group X^2	N	P
1	Trash	75.03	80	.000**	12.5	21	.000**
2	Trail	6.05	81	.014**	Binomial*	20	.031**
3	Trail	0.00	79	1.00+	Binomial*	20	.125+
4	Trash	72.01	80	.000**	10.56	20	.001**

*indicates X^2 coefficient not computable because 1 x 2 table

[1]based on McNemar Test

**significant

+recalculated value : experimental group X^2= 49.4, n=81, P=.000
control group X^2= 11.1, n=20, P=.001

response had to be chosen as "correct" and all others had to be disregarded. For question three, the response, "run around a short path that takes you to the plant" was chosen because it was closest to program objective 11. However, the majority of individuals in both groups chose "ask the ranger if you'll see other plants like it later on" as their response, and this response was also not incorrect. When the statistic was recalculated using the second response, "ask the ranger...plants like it later on", the change was significant.

As with littering behavior, then, behavioral intention responses were essentially the opposite of the actual behavior. These data seem to suggest that the program effected participants in an undesirable way. However, this result must be cautiously interpreted because the methods may not be truly comparable.

These data also suggest that both staying on the trail and littering behavior need to be measured using observation rather than behavior intention instruments in order to get an accurate picture of behavior. Yet one must carefully apply this interpretation given the teacher's influence on the control group's behavior and the small sample size from which the results are generated.

Discussion: Study Objective 1 (Evaluation of Program Effectiveness)

Knowledge Results

Since the program objectives were not met, few statistical changes were found and only certain concepts seem to

stand out as important knowledge information, it may be helpful to speculate why the results occurred in the manner in which they did.

First, the instruments themselves may not have accurately measured the individuals knowledge (or lack of knowledge). The instruments did not appear to be reliable. Therefore, the changes that resulted may be due to the instruments rather than the treatment effects. Also, the instruments would not have given accurate results about the program's effectiveness. This suggests that future researchers should develop instruments that are known to be reliable. The responses which were obtained for the levels of schematic differentiation also suggest this might be the case. One would expect that the more complex the question, the fewer percentage that would be answered "correctly." In many cases just the opposite occurred, i.e., the greater precentage of individuals answered the more differentiated questions "correctly." In this study, since only five levels of differentiation were used (Table 20, Figure 4), and between one to three questions were used to measure a particular differentiation level, it is difficult to say whether this pattern is an accurate measure of the individuals' schematic organization. In addition, only a small sample size made up the entire study, which further limits what one may say about individuals' schematic organization.

Another indication the instruments may not have accurately measured individuals' knowledge is found in the number of

"partially correct" answers that were given (Table 24). For the crossword puzzle, "partially correct" was defined as the correct number of letters, but the wrong word in the answer blank (they were not awarded points for these answers when their scores were calculated). Before the program, the experimental group averaged 16.9% "partially correct" and the control group averaged 13.1% for all knowledge questions (Table 24). Similar results were found with the posttest: the experimental group averaged 7.4% "partially correct" and the control group averaged 6.4% for all knowledge questions (Table 24). The researcher observed at least one-third to one-half of students from every school counting the number of boxes in the crossword puzzle and then trying to fit a word from the spelling list into the boxes, suggesting that students were guessing. Hence, a realistic estimate of their knowledge would not have been achieved.

For the matching game, the correct answer was four boxes (for seven out of eleven questions), but three, two or one constituted a "partially correct" answer. The experimental individuals who completed the matching averaged 55.8% "partially correct" and on the posttest they averaged 33.2% "partially correct" for all knowledge questions (Table 24). In other words, the individuals got a greater percentage of partially correct before they were exposed to the treatment than after. The percentage of correct answers remained approximately the same before and after the treatment. It seems that students might have been guessing on

Table 24. Percentage of Partially Correct for the Knowledge Instruments for the Control and Experimental Group

Program Objective Number (pretest; posttest)	Program Objective Content	Experimental Group Pretest: % of Group Partially Correct	Experimental Group Posttest: % of Group Partially Correct	Control Group* Pretest: % of Group Partially Correct	Control Group* Posttest: % of Group Partially Correct
			Crossword Puzzle		
1,2 (1,3,6,11); (9,5,4,2)	know plants prevent erosion, know walking on plants kills them	9.3	1.7	1.2	4.2
3 (5,9); (6,11)	know plants/ animals are protected	29.4	10.0	21.4	8.3
4 (2,4,8); (7,3,1)	know animals have specific habitats	12.5	12.2	9.5	6.9
5 (7);(5)	know a successional community	21.6	8.3	23.8	4.2
general (10);(10)	-	11.8	5.0	9.5	8.3
			Matching Game*		
1,2 (6,7,8,3); (8,2,9,4)	know plants prevent erosion, know walking on plants kills them	63.1	48.9		
3 (2,10); (6,1)	know plants/ animals are protected	78.3	21.7		
4 (9,4,1); (10,7,3)	know animals have specific habitats	50.7	21.7		
5 (11);(11)	know a successional community	0.0	0.0		
general (5);(5)	-	87.0	73.9		

*A control group did not fill out a matching game

this instrument (especially pretest) because a greater percentage answered "partially correct" and they finished it very quickly (approximately two-thirds finished it in about five minutes). When they did the posttest they took longer. In fact, some felt very frustrated when personnel gathered the instruments. When individuals were doing the posttest, they might have thought more about which answer was correct as they had just heard the answers plus other relevant information. Also, when Grand Ledge students pilot tested the instruments, they completed the matching very quickly and told the researcher they felt that they were guessing.

The differences in the percentages "correct" for the two different knowledge test types may have been due to the nature of the tests. The crossword puzzle, which included spelling words, essentially measured visual recall and hence may have been easier to complete. The matching test, on the other hand, gave no "clues" about the "correct" answer and was therefore more difficult. An analogy may help illustrate this point. A true-false, matching, or multiple choice test is like the crossword puzzle where the possible answers are given and the respondent chooses one of the answers. An essay test is similar to the matching game, where the respondent must answer based on his/her own knowledge and is not given any hints from the examiner. Usually, essay tests are considered more difficult but more suitable measures of knowledge. In fact, education specialists such as Hoffman (1967) have criticized multiple choice tests

and have recommended using essay tests in order to better measure subjects' knowledge. When one views the knowledge tests from this perspective, the matching game may have been a more accurate test of participants' knowledge than the crossword puzzle.

Second, teachers may have exposed students in both control and experimental groups to the concepts that they would be learning on their field trip to INDU. For instance, at Edgewood School, some of the students used a ditto handout to complete their knowledge test. As soon as the researcher realized they were doing this, she had them put the dittos away, but in many cases they had already used them. In the future, park personnel should contact teachers, explain the purpose of the research and ask the teachers to cooperate in the study (for example, do not expose students to program content information).

Third, over one-third (35.7%) said they had visited INDU before. One-fifth (20.2%) said they visited with school, 21.4% indicated they had visited with their family, and 15.5% said they had visited with both family and school. It is possible that on their previous trip(s), particularly with school groups, they learned sufficient information so their pretest responses did not statistically change from their posttest responses.

Fourth, the learning process is very complex and often poorly understood. For example, people selectively perceive and recall information, i.e., they filter out information.

They do this for several reasons. People may not understand the information and hence cannot assimilate it. They may be overwhelmed by the amount of information. The information may not be interesting to them. The information may contradict existing values, attitudes, beliefs and behavior which are more ingrained and central to the person than the new knowledge.

Because learning phenomena such as selective perception and recall exist, it may be difficult for any individual to gain a statistically significant amount of new information during a one and one-half to two hour program. But some incremental knowledge gain seems possible. As stated previously, the largest knowledge changes were knowing what a dune was and knowing that plants and animals should be protected.

Fifth, the design of the program might help explain the results. For instance, the program consisted of many isolated facts that were not tied together by a common theme. The program lacked unity. Many of the interpreters used jargon rather than descriptive language. For the most part the programs were not designed to encourage involvement. The program (treatment) did not follow many other principles of "good" interpretation. (Table 1 lists principles that are needed for "good" interpretation.)

Theorists and researchers in education and communication such as Piaget, Gagne and Bruner have suggested certain principles that should be followed in order to facilitate

learning. Interpretation specialists have also recommended specific principles that need to be followed for effective learning to take place (see for example, Wagar, 1977). In interpretation, for example, the message should have an underlying theme which is the main concept of the program. This theme should be clearly apparent and woven throughout the program. Although many of the established principles were not utilized in this program, these principles have not necessarily been proven to be effective themselves (Fritschen, 1980). Yet, to date they are the best guidelines available, and therefore should be used in interpretive programs.

Sixth, a common problem with evaluation studies lies with the objectives themselves (Suchman, 1967; Theobald, 1979). The percentages specified in the program objectives were not met for nearly all objectives for both the control and experimental group posttests. This result is expected for the control group as they were not exposed to the treatment. However, the experimental group should have come closer to meeting the objectives. This indicates that the objectives may have been set too high (or the program did not impact the individuals or the instruments were problematic, etc.; see previous discussion). Program evaluators (Johnson, personal communication, 1983; Rossi, 1972) have recommended that objectives should specify success rate around 50% Almost all the program objectives for this study specified 80% to 90% success rate and only two out of twelve program

objectives specified a success rate of 50% or below. If they had been set around 50%, most all knowledge program objectives would have been met for those who filled out the crossword puzzle, and for those who completed the matching objective three (protecting plants and animals) would have been attained. However, setting the objectives around 50% would introduce additional problems as indicated by the results of this study. The (approximate) 50% "correct" criteria was met on the pretest for experimental group who filled out the crossword puzzle and the control group. The 50% criteria was also met for most objectives for the posttest control. The matching percentages "correct" before were also similar to those after the treatment. Because these problems exist, setting the objectives around 80% to 90% in this case may have been appropriate. These success rates, however, will probably not be achieved with a single visit (in future). As the objective success rates were not achieved and setting them at either 80 to 90% or 50% creates several problems, it may be more suitable to try to examine changes that may have resulted from the program.

Often the objectives are not adequately grounded in theory and/or they are not based on any particular truth or concept (Suchman, 1967). This may have been the case with this study as INDU personnel compiled the program objectives based on their perceptions of what participants should know (feel and behave). It appeared that the objectives were written in a haphazard manner and little thought went into

the reasons for using each objective to begin with. This may be determined by closely looking at each of the knowledge objectives. The program objectives were supposed to be behavioral objectives that were measurable. But program objectives one to three say "participants will know..." Knowing is not a behavioral objective term. Words such as "verbally defines," "describes," "names," "matches," etc., specifically indicate what behaviors are desired. In fact, Bloom (1956) has proposed six levels of behavioral objectives for measuring the cognitive domain, and corresponding verbs that describe these levels. These verbs should be used in future studies. The objectives are vague in other ways as well. Program objective one says "marram grass and other plants" but the other plants were not defined. Program objective four says a "common dune animal and plant," however, again these were not specified. Although this information was eventually obtained from INDU personnel, the ambiguity of the program objectives suggests that personnel put them together in a rather unorganized way.

Seventh, though content was kept as constant as possible, by using a content checklist, the content varied somewhat in each program. For instance, an interpreter did not mention the bearberry bush and raccoon to Edgewood School individuals (n=27), yet questions about these organisms were on the matching posttest. If the information was not covered, it would be difficult for individuals to learn it. However, a careful and complete review of the content checklists

showed the content was similar (as much as humanly possible) for each program.

Attitude Results

Because many factors influence attitudes and attitude formation (Shafer and Tait, 1981; Robinson and Shaver, 1973), respondents' attitudes may not significantly shift after they have been exposed to a two hour program. For example, the family and television (media) greatly influence people's attitudes. Their previous experience, cultural background, social mores, etc., also influence attitudes and attitude formation.

Attitudes are highly interrelated (Shafer and Tait, 1981; Peyton, 1983) and in order for one attitude to change or shift it may be necessary that several others change or shift also (Shafer and Tait, 1981; Peyton, 1983). The attitudes that were measured in this study (except protecting plants and animals) may be more central than peripheral to the individuals and hence will not shift as a result of being exposed to a two hour program.

The instruments may not have accurately measured participants' attitudes. They were not highly reliable (Table 13). The questionnaire did not measure all the attitudes that an individual can have about any one particular object, much less general objects such as plants and animals. Hence, the individuals' attitudes may have shifted or changed on a micro-level but the study's instruments would not have measured these micro-level shifts. The instruments were designed to measure attitude specific information about INDU, but

participants might not have known what INDU was and/or had any feelings about it.

The program content did not specifically address appropriate attitudes (the instruments' questions were based on the objectives, not content, although, the content is reflected in the program objective) that the individuals should have as a result of the program. Since this information was not overtly given, and it is generally agreed that attitudes are a learned phenomena (Fishbein and Ajzen, 1975; Kaplan, personal communication, 1984), the individuals probably kept attitudes that they had (always) maintained in the past. If the program content had addressed certain attitude orientations (based on the objectives, such as attitudes about littering, protecting plants and animals, etc.) it is doubtful that this information would have significantly affected the individuals as they had the desired attitudes even before the treatment (see "Attitude: Pretest Responses" section).

Behavior Results

Behavior itself as well as behavioral intention are very complex. Factors such as previous behavior, education, media influence, existing attitudes, cultural background, opportunity, and folkways influence and motivate behavior and intention (Fishbein and Ajzen, 1975). Because they are so complex, it may be difficult to change behavior and intention from simply exposing individuals to an interpretive program.

The program's content was not geared toward behavior and behavior change, rather the majority of information focused on ecological facts and concepts. While knowing these facts and concepts may in some way be related to the behavior that an individual performs, the relationship is not direct. In other words, participants can know and feel a desired way about a certain behavior but not actually behave that way. Since the relationship between behavior and knowledge (and attitudes and values) is so complex, it is necessary for INDU personnel to decide what domain they really want to affect and concentrate their program efforts on the particular domain.

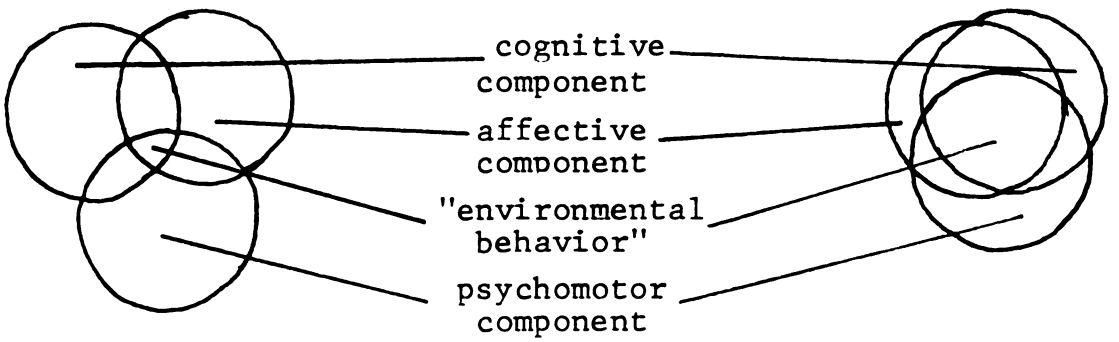
The pretest behavior instrument may account for the deviation between behavior intention and behavior which resulted. Question one may have been leading because it said "drop it in the trash can," yet all other possible responses, except clearly the incorrect answer began with the verb "put." In future, additional care should be taken to make sure the "correct" answer does not tend to stand out. For example, start all responses with "put." Questions two and three contained more than one "correct" answer. Although only one answer was "correct" according to the objectives, other answers were not really incorrect, such as "ask the ranger if you'll see..." To correct this problem, a single "correct" answer based on the objectives should be provided on future instruments and the other responses should be reasonable (from the respondents viewpoint) but not "correct."

If this procedure is followed, the results will give a more accurate idea of their behavior intention because it will be easier to discern between right and wrong answers. Also, questions two and three asked respondents how they would behave on or near a trail if they saw an interesting plant or animal. While this story may give some indication of their trail behavior, other stories may be even better. For instance, participants previously said they liked running down hills or dunes. A story that asked participants if they would run down a vegetated dune or stay on a flat trail might give a better indication of their behavior.

The deviation between the actual behavior and behavior intention may have resulted for other reasons. The measurement methods were different. This suggests that future studies should be designed so that the method for measuring pretest and posttest behavior is the same. The trash can was not conveniently located for all groups, so it may have been easier for the control group to properly dispose of their trash than the experimental group.

Study Objective 2: Evaluation of Interpretive Methods

The trait "environmental behavior" was measured using four methods (game, attitude questionnaire, behavior intention stories, observation). It was operationalized as having cognitive, affective and psychomotor (behavior) components (Figure 14). The essence or content of each component of the "environmental behavior" trait was derived from the program objectives. For instance, the cognitive



a. weakly correlated b. strongly correlated

Figure 14. Components of the Environmental Behavior Trait.

component consisted of knowledge about protecting plants and animals, plants preventing erosion, a dune, and a successional community.

If the methods used to measure the components of the "environmental behavior" trait were strongly correlated, then one method of measuring the trait may be substituted for another method (validity correlation) (Figure 14b). The validity correlation(s) (same trait, different methods in this study) should be smaller (approximately .5 - .7) than reliability correlation(s) (same trait, same method) (approximately .7 - .9) and larger than nonsense correlation(s) (different traits, different methods) (approximately .1 - .4) (Kidder, 1981).

Although some of the methods (questionnaire, paper and pencil games, observation and behavior intention story) were significantly correlated they were not found to be strongly correlated in terms of magnitude. (Statistical significance is not meaningful in correlation analysis

because large sample sizes (n>30) are usually significantly correlated although they are not necessarily strongly correlated.) The positive correlations between the methods ranged from .013 (behavior intention stories and pretest questionnaire) to .440 (posttest questionnaire and posttest crossword). The negative correlations between the methods ranged from -.001 (behavior intention story and posttest questionnaire) to -.188 (behavior intention story and posttest crossword) (Table 25). These results indicate that the methods cannot be substituted for each other, i.e., to get a complete portrayal of the "environmental behavior" trait (as defined for this study) it is necessary to use all four methods.

This seems logical for several reasons. The methods' test items were not reliable (Tables 12,13). Also, all the knowledge instruments (crossword puzzle, and matching, pretest and posttest) were more strongly correlated to the questionnaire than themselves. In other words, the correlation between the same method (knowledge instrument) which measured the "environmental behavior" trait (i.e., reliability correlation) was smaller than the correlation between two methods (questionnaire and knowledge instruments (i.e., validity correlation). The opposite results were expected. Hence, future research must be based on instruments that are known to be reliable. For example, correlation analysis should be performed before the instruments are used in the field. A panel of experts should evaluate the instruments. Previously developed instruments should be located and used

Table 25. Correlations Between Methods[1]

	Method						
	Questionnaire		Game				Observation
	pretest	posttest	Pretest matching	Posttest matching	Pretest crossword	Posttest crossword	
Experimental Group							
Method: Game	.100(23)*						
pretest matching	.342						
		.351(22)					
posttest matching		.054					
	.208(51)						
pretest crossword	.072						
		.276(60)					
posttest crossword		.016					
		-.307(21)				-.099(15)	
Observation		.088				.363	
			.167(25)		.372(51)		
Behavior Intention			.220		.010		
Story	.193(82)						
	.041						
Control Group							
Questionnaire pretest		.531(22)					
		.006					
posttest							
Method: Game	.379(21)	.346(21)				-.025(21)	
pretest crossword	.045	.062				.457	
	.301(22)	.440(22)					
posttest crossword	.087	.016					
	not computable	not computable			not computable	not computable	
Observation	(4)	(6)			(4)	(6)	
	.013(22)	-.001(22)			.086(21)	-.188(22)	not computable
Behavior Intention Story	.447	.498			.351	.201	
Control and Experimental Group							
Game	.235(76)						
pretest crossword	.020						
posttest crossword							
Method: Observation							
	.139(104)				.253(76)		
Behavior Intention Story	.079				.014		

[1] correlations for n < 6 were not reported
*correlation coefficient (n) significance level

if possible (see Journal of Educational Measurement, Journal of Educational and Psychological Measurement, Evaluation in Education and others).

In addition, "environmental behavior" was very broadly defined. It included content information about knowledge, attitudes and behaviors that were not known to be related. In psychological testing, however, the multitrait-multimethod correlation analysis has been used on content that is more defined (depression for example). The content included in the methods used to measure the trait is known to be related to the trait. (For instance, people who are depressed behave in certain ways, and tend to have certain attitudes. Researchers might use observation to watch these behaviors and perhaps a questionnaire or interview to find out about their attitudes. The two methods would then be correlated to determine if they could be used interchangeably.) Therefore, in future research in interpretation more specific trait(s) needs to be investigated. For example, a littering trait, protective trait or "staying-on-the-trail" trait are more specific than an environmental behavior trait and should be examined in future studies.

CONCLUSION

Summary

Study Methods

This study was designed to accomplish two objectives. Study objective one was an evaluation of interpretive program effectiveness at Indiana Dunes National Lakeshore. Effectiveness was operationalized in three ways: 1) extent to which prestipulated program objectives were met; 2) changes in knowledge, attitudes and behavior as a result of the program; and 3) changes in participants' level of schematic differentiation as a result of the program. One hundred-twelve fourth to sixth grade individuals from schools nearby the Lakeshore participated in a pretest-posttest control group study design. Four methods were used to evaluate the program: paper and pencil games (crossword puzzle and matching game) to measure knowledge, questionnaires to measure attitudes, behavior intention stories to measure behavior before the program and observation to measure behavior after the program.

The methods used for study objective 1 were investigated for study objective 2. For this objective the various methods were correlated together to determine if the methods could be used interchangeably in the interpretive setting to measure an "environmental behavior" trait (as defined by the study objectives).

Achievement of External Criteria and Treatment Effects

The prestipulated program objectives (external criteria) (pages 33-34) were not attained for either pretest or posttest knowledge and attitude measurements and not for posttest behavior measurements. The prestipulated criteria were achieved for pretest behavior measurements. Although the external criteria were generally not satisfied, some treatment effects were realized.

Knowledge Results

For the most part, the experimental group "scored higher" than the control group on the posttest and the experimental group's responses were "higher" on the posttest than on the pretest. In particular, their responses changed statistically significantly for protecting plants and animals and changed somewhat for knowing what a dune was and knowing that plants hold the dune in place to prevent erosion and that walking on plants can kill them which in turn causes erosion.

Attitude Results

The majority of both the control and experimental group had positive attitudes about protecting plants and animals and negative attitudes about littering both before and after the treatment (program). Both groups also felt positive about nature parks in general and over 90% wanted to return to INDU. In addition, the majority indicated they would talk to their parents (families) and friends about their field trip. Most of the respondents in both groups indicated (on both pre- and posttests) that they would most like

to run on hills without plants. In fact, when asked what they liked best about their trip, several said they like running on the hills best. Over 80% of the experimental group said they enjoyed the program or enjoyed it very much.

Behavior Results

For staying on the trail and littering-disposal behavior, respondents in both the control and experimental groups indicated that they intended to stay on a marked trail and put their trash in a trash can, respectively. The majority of participants (both groups) went on an unmarked trail. Most students in experimental groups (66%) put their trash in their pockets and 19% dropped it on the ground. All control respondents put their wrapper in the trash can.

Discussion

The pretest data suggest that respondents had desirable attitudes and were generally aware of (some) park rules and suitable park behavior. In most cases they were more aware of rules and appropriate behavior than certain ecological facts or concepts. The posttest data suggest that more experimental individuals were knowledgeable about a dune after the program. More experimental respondents knew about and had stronger attitudes about protecting plants and animals. Respondents also had stronger anti-littering attitudes after the program. The posttest behavior data suggest that respondents did not behave suitably.

The trends found in the data must, however, be carefully interpreted because many factors may have influenced them. Many items on the instruments were found to be

unreliable, suggesting that the changes may be due to the instruments rather than the program effects. Over one-third of the individuals had previously visited the park; these previous trips may have influenced their responses. The control group's teacher told the control group how to behave (during posttest behavior measurements). Also, the influence that teachers had on the students in the classroom was not accounted for. The majority of the program focused on knowledge, yet behavior and attitudes, as well as knowledge, were measured to determine the program's effectiveness. Finally, numerous unmeasured factors influence knowledge, attitudes, behavior and intention. These factors may have effected the results.

Measurement of Schematic Differentiation

The percentage of participants' correct responses varied for each level of schematic differentation. The largest percentage of correct responses (pretest and posttest; control and experimental groups; crossword puzzle and matching game instruments) were found for levels two and three. Both these levels had questions which solicited participants knowledge about behaviors. This trend suggests that participants' high knowledge of behavior skewed the response pattern.

Evaluation of Methods

Knowledge games, attitude questionnaires, behavior intention stories and observation methods were weakly correlated to each other. These data suggest that all four methods are needed to establish an accurate description of

the "environmental behavior" trait (defined by the study's program objectives).

Implications and Recommendations

The combined results from study objective one and two have practical and theoretical implications. Recommendations and future research directions based on these implications are listed below.

Practical Recommendations

1. Emphasize resource management practices and the associated proper behaviors for the dune environment in future interpretive programs. Use ecological concepts and facts as background material.

2. Gear interpretive programs to changing behavior. Even though participants knew certain "correct" behaviors and had knowledge and attitudes which seemed consistent with the behaviors, the behaviors did not actually occur.

3. Set program objectives around 80% success rate but do not expect to meet them without repeated programs for the same individuals. OR Investigate changes that may have occurred before and after the program. (For example, observe participants before and after programs to see if their behavior has changed.)

4. Write specific, concrete, descriptive and measurable objectives to use as a basis for evaluating and guiding future programs.

5. Implement management practices that meet visitor needs and make certain all aspects of the park (from personnel to resources to rules) create positive feelings (and images) for visitors. To make sure this task is accomplished, it will be necessary to find out about park visitors: their needs, likes, behaviors, backgrounds, etc.

6. Designate a dune for climbing and running on to meet visitor needs and protect the park's natural resources.

7. Conduct programs that instill positive feelings about parks. This recommendation implies that park personnel must know what creates/facilitates positive feelings in visitors so these activities, etc., can be incorporated into the program. For instance, if visitors feel running down dunes is fun and they want to do it, but this activity is not included in the program, visitors may develop negative attitudes about the interpretive programs and the park and hence not return.

8. Design trails that are practical from the visitors' perspective. In other words, determine the most common visitor travel pattern and use these patterns as a basis for the trails. Clearly mark these trails so visitors are not tempted to start a trail of their own. Make certain the trails are straight paths, as visitors will tend to take the shortest distance between two points because it is easiest.

Evaluation Research Directions

1. Investigate more fully the methods used to evaluate interpretive effectiveness. For example, use the multi-trait multimethod matrix on a more defined trait, such as "litter-disposal behavior," "staying-on-the-trail behavior," or "protection behavior." Design the instruments so the content is related to the trait (for instance, develop a questionnaire that solicits "protection" attitudes).

   Examine both discriminant and convergent validity. Convergent validity was the only constructive validity type that was investigated in this study.

2. Take two new approaches to try to discern more clearly what effect the programs have on participants. First, include program information which may be measured with "a fine tooth comb." In other words, investigate more completely the more subtle effects of the program. Second, devise more diverse program objectives. Differences in attitudes, knowledge and behaviors that were not measured in this program may have occurred and it would be helpful to know what these might be.

   Both of these approaches may require using a different audience, mode of interpretation or adding certain audio-visual devices, etc. These possible program changes should be pilot tested before they are put to use in a study environment.

   Both approaches would also require that the instruments themselves be revised. This task includes checking for and then using instruments that have already been used for measuring various attitudes, knowledge and behavior. For example, Miller (1964), Shaw and Wright (1967), and Robinson and Shaver (1972) have designed various attitudinal scales that may be useful for future studies. Numerous journals and books, such as Educational Measurement, Educational and Psychological Measurement, Educational Psychology, Evaluation in Education and Constructing Evaluation Instruments (Furst, E.J. 1958; New York: Longmans, Green and Co.), may be useful.

Interpretive Evaluation Instruments: Recommendations

1. Check all instruments for reliability and validity. To implement this recommendation, research might be conducted to develop evaluation instruments which can be used for numerous agencies' and organizations' interpretive evaluation.

2. Use a matching game or similar instrument to measure knowledge. It appears to give a more accurate measure. It seems to test more what participants know since the instrument does not give any answer clues like the crossword puzzle does. Because the matching game is designed with boxes, where one, two, three or four may be filled in for the correct answer, it is also a better measure of levels of schematic (or cognitive) differentiation.

3. Use the same number of concepts for each schematic level and try to use more than five levels to see if a pattern emerges, if the schematic differentiation approach is employed. Use only ecological concepts to measure levels of schematic differentiation (or only behaviors).

4. If behavior intention stories are used, design the instruments so that only one response is truly correct and the other responses are reasonable to see if respondents actually intend to perform the correct behavior.

Limitations

Although the study was designed to control for as many variables as possible, the data must be viewed with some caution. Various limitations have resulted from the practical, financial and temporal restrictions. These are discussed below:

General Limitations

1. Values were not included in this study, yet values often may be weighted more heavily in decisions (for example, about behavior) than knowledge and attitudes.

2. Many, many factors affect each behavioral domain (cognitive, affective and psychomotor). For example, one may feel it is important to protect plants and animals, but at the same time he/she may want to go back to the park to ride dune buggies.

3. It is unknown what specific content in a program teaches what specific behavior, knowledge or attitude.

4. Some teachers prepare students for the trip before they go to the park, which might cause individuals to have higher pretest and posttest scores.

5. Some students have previously visited the park which might affect their responses.

6. The results may not be generalizable to a population of fourth to sixth graders for several reasons. Only age and gender demographics were collected from the individuals so it was not possible to compare demographic information with other schools. The sample size was quite small (due to factors which the researcher could not control), and not truly random. The teachers' motivation to bring students to the park largely influences whether the students come (Vaughn, personal communication, 1983). Chances are the teachers who bring their students to programs are also more keen to expose their students to environmental information and issues than teachers who do not bring their students to the park. If this is the case, the students already have more background and hence are not "random" individuals.

   Time constraints prohibited the researcher from comparing the sample schools with the area schools to see if they appeared to be representative, regardless of the teachers' motivation.

Study Limitations

1. The amount of information that may be elicited from children is limited.

2. It is logistically difficult to administer paper and pencil tests in an outdoor setting.

3. The observation method was not checked for reliability because staff were limited. Ideally at least two staff should be available in addition to the interpreters to help with the observation.

4. The methods used to measure behavior before and after the program were not truly comparable. Ideally two additional staff are necessary. One could administer the posttests to the control group and watch their behavior. The second person could observe the experimental group before they participated in the program. This solution still had problems, however, because the control individuals are not observed twice.

5. The excitement of going to the park may cause change in the responses (even for the control group). Therefore, it may be better to test people who do not go to the park at all.

6. The differences found between the control group and the experimental groups' posttests and the pretest and posttests for the experimental group are quite small. They may be due to chance rather than a program effect, especially considering the differences are based on a small sample size and somewhat unreliable instruments.

Limitations of Study Site

1. The experimental main trail was poorly marked and difficult to distinguish from other unmarked trails.

2. Few areas were really suitable for observing behavior.

3. Congestion occurred when more than one school group was on the trail at the same time.

Limitations of Instruments

Although the limitations of the instruments were accounted for as much as possible, they should also be acknowledged.

1. Although the instruments were tested for reliability by pilot testing and the wording in the before and after instruments was very similar, reliability correlations suggested they were not reliable. (These results, however, are based on a small sample.) Ideally, correlations should have been run on the pilot test individuals, however time and financial constraints prohibited this. Also, if time and money permitted, a panel of experts should have rated the equivalent questions to determine if they meant the same thing. Revisions would then have been made based on correlation analysis and the panel's comments.

2. Other questions which measure these program objectives may already have been developed. In the ideal situation, the researcher would look for questions that were already designed (and presumably tested for reliability and validity) to measure the concepts stated in the program objectives.

In sum, the study results indicate that many more questions about evaluating interpretive effectiveness and effectiveness measurement methods must be answered. As a first step to exploring these questions, additional data for this study (collected in the spring of 1984) will be analyzed and available in a management report.

APPENDIX A: INVITATION TO PARTICIPATE LETTER AND WEST BEACH BACKGROUND INFORMATION

TEACHER REGISTRATION FORM

Name_____ Program Date & Time_____

School_____ Place(s) of Program_____

Address_____ _____

_____ Number of Students_____

Grade Level of Students_____ Number of Teachers_____

Number of Other Adults_____

1. Will this be the first trip to Indiana Dunes National Lakeshore for this group of students?

_____ Yes

_____ No

If not, what areas have you visited on previous field trips?

_____ West Beach _____ Mt. Baldy

_____ Bailly Area _____ Ly-co-ki-we Trail

_____ Cowles Bog _____ Other (Please specify) _____

2. Please estimate how many of your students have visited Indiana Dunes before.

_____ (with a class) _____ (with family, friends, etc.)

3. Have you personally, ever been to Indiana Dunes National Lakeshore before?

_____ Yes

_____ No

What areas have you visited?

_____ West Beach _____ Mt. Baldy

_____ Bailly Area _____ Ly-co-ki-we Trail

_____ Cowles Bog _____ Other (Please specify) _____

4. Please list one or two reasons why you are bringing your class on this field trip?

5. Will this trip initiate, continue, or complete a unit of study? Please circle one in the previous sentence, if applicable.

   If so, briefly describe the study.

   Are there specific ways that we could relate the field trip to your study?

6. Please list one or two activities which you plan to do in the classroom to prepare your students for the field trip.

7. List any medical conditions or disabilities you are aware of (i.e. all heart conditions, hypoglycemia, epilepsy, etc.) that may effect the vigor of anyone participating in this program.

Thank you for your help! By signing and returning this form you will be able to follow the guidelines listed on the preceeding pages.

_____

Signature

_____

Date

TRAIL DESCRIPTION WEST BEACH

West Beach is a unique and beautiful part of Indiana Dunes National Lakeshore. Classical sequences of dune succession appear in this area. Here you can find some of the dunes' secrets: delicate interdunal ponds nestled behind the first dune ridge; moving or living dunes gradually creep inland; many species of plants and animals are unique to this environment.

Most groups visiting the West Beach area are interested in dune formation, dune succession, and human impact on the dunes environment, so these are the topics emphasized in our field trips. At the same time, we also like to stress the development of an increased environmental awareness. It is important to us that students learn to better appreciate the natural world around them and how to use the park without destroying it. To this end they will not only be expected to follow certain rules, but will be involved in many learning activities. These activities will include both explanations of the processes at work and participatory games, such as role playing and blindfold walks. In many of these activities participants will be encouraged to use their senses. Please remember that it is imperitive that the teacher, and other adults accompanying the group be full participants to all of these activities. Students need your example, especially when the class is removed from the familiar classroom environment.

When you come to West Beach be prepared to hike one to two miles over hilly terrain and loose sand. Several trails are available in the area, and your hike may take you along the

beach, through the dunes, or into the woods, depending upon the age and interest of your group as well as the amount of time you have to spend at West Beach.

What you, your students, and your chaperones wear is very important. Platform shoes, high heels, and sandals are dangerous and uncomfortable, and should not be worn. Tennis shoes, hiking boots, or sturdy walking shoes are suggested. Jackets or raincoats are necessary if it is cool or cloudy. In mosquito and deerfly season we strongly recommend insect repellent.

During the summer months West Beach is the main swimming beach area for the National Lakeshore. At that time only a "user's fee" of one dollar per car is charged for parking in the area. In summer the bathhouse is open and a concession stand, changing facilities, and showers are available. Lifeguards are also on duty during this period. There are picnic areas and tables in the West Beach area. You can, also, picnic on the beach or in the dunes if you pick up your trash. Glass containers are forbidden. Drinking water is available only when the bathhouse is open (approximately Memorial Day through the first fall frost). Restrooms or chemical toilets are available all year.

APPENDIX B: PROGRAM CONTENT

WEST BEACH PROGRAM DESIGN

THEME: Adaptation and human impact

LOCATION OF PROGRAM: West Beach

INTENDED AUDIENCES: 4th through 6th grade

Sequence of Stops

STOP 1 (Sidewalk)

1) Introduction is done at the West Beach bus parking lot. Welcome the group and introduce yourself. During the first part of the introduction, point out the following):
2) This is a national park and it is called the Indiana Dunes National Lakeshore.
3) Question: Can you name some other national parks? (get random answers)
4) Question: Has anyone visited some of these other national parks? (get random answers)
5) (Point out similarities and differences between the National Lakeshore and some of the other national parks.):
   a) Some national parks are large enough that one can go hiking or backpacking for quite sometime without seeing another person.

b) This Park is fairly small, and is surrounded by industries, houses and highways. It is especially important that visitors treat it with care.

c) This Park has many beautiful and fragile areas: West Beach is one such area.

6) This is the West Beach area, one of several areas in the National Lakeshore. (Using a map, locate West Beach and a few other INDU areas).

7) West Beach is the most popular recreational area of the National Lakeshore because of its beach-house and swimming area. However, there are many other interesting things to discover at West Beach -- hiking trails, ponds, flowers, wildlife; we will see and discuss many of these later. But these plants and animals are threatened by the large numbers of people using this area. We will see what we need to do to protect them.

8) We will be hiking to the beach (point in that direction). There we'll go back in time and discuss the history of Lake Michigan and its sand dunes.

9) Then we will hike east on the beach (point in that general direction) and explore the shoreline and beach, and see how dunes are formed.

10) We will then hike south in the dunes (point in that general direction). During this time we will:
    a) Discover different "neighborhoods," or "communities" of plants and animals. (succession)
    b) Visit an interdunal pond, a dune blowout and an oak forest.
    c) Look for signs of wildlife.
11) We will be discovering ways that plants and animals solve the problems of staying alive, called adaptation. (have them repeat the word). Hike back to the bus in about 2 hours.
12) (Discuss rules of the hike):
    a) Staying on the trail
    b) Enjoy flowers and other plants without picking them
    c) No littering
    d) We'll watch animals from a distance, and let them live their lives.
    e) I would also like to have your full attention and participation during the program (agree on specific standards, such as "we won't talk while another person has the floor").

    Activity:

    group contract

STOP 2 (North Parking Lot Plaza)

Activity:

Snake & turtle

(Have everyone look around from where they are standing and see if they notice the different sizes and shapes of sand dunes around them.)

At one time there were very tall sand dunes right where we are now standing. Many sand dunes were removed from this spot many years ago - sand was used to make glass, sand paper, landfill, cement, etc. This is called sand mining. Therefore, when the National Park Service decided to build a beach house, this damaged area was chosen as the parking lot site. (Point out Visitor Center, offer a bathroom stop, then continue hike.)

STOP 3 (Beach house road)

(As you walk along the road to the Beach House, point out the erosion, trails and blowout on the left. Introduce the problem of trail erosion. Tell them that we're marking certain trails for hiking, and closing others. Ask them for the cooperation).

STOP 4 (Beach House)

Question: Do you notice anything unusual about this building? (Get random answers. Have group notice the way the building is elevated on stilts, so sand can blow underneath it.) You may have

also noticed that this building is hidden by the dunes surrounding it. (Remind group to stay on the trail. Mention and briefly observe pond: more later.)

STOP 5 (Beach - just east of the bath house)

Deemphasize interaction in lecture groups (Allow for excitement time; participants should have the opportunity to look around on their own for a few minutes. Regroup and form semi-circle and briefly discuss the following):

1) The Wisconsin Glacier
   a) Huge mass of ice, snow and earth debris, one mile thick, that moved slowly.
   b) Came from the upper regions of Canada, (point in general direction) and covered northern Indiana and surrounding areas.
   c) A rise in temperature caused the ice to melt.
   d) When the glacier melted, Lake Michigan was much bigger than it is today. (Point out to group that their school and homes may be located in an area once covered with water.)
2) Lake level soon began dropping. Its present level is 500 feet above sea level, 60 feet below what it was when the glacier melted.
3) Several shorelines were formed as lake level dropped. Geologists have identified three

major old shorelines. You can find these shorelines today. (Relate to the locations of their schools and homes.)

4) Between each dune ridge you will notice, in some areas, marshes or swamps. Very few have been left: most have been drained. They are rich in wildlife, important for flood and pollution control, and should be protected. Notice Long Lake and the Great Marsh from your bus as you leave.

STOP 6 (Just east of the last stop on the beach)

1) Question: Where does all of this sand come from? (Get their responses: go through #3 before discussing "right" answer).

Activity:

soil sample #1

Activity:

sand through the lens

Over thousands of years storm waves have eaten away at the cliffs along the Wisconsin and Michigan shorelines. Waves and currents carry on the work of ancient glaciers, and act as giant grinders as stones are broken down to form sand. Streams wash more sand into the lake. The sand is washed up onto the shore, it dries, and the wind blows it inland. Sand piles up around the dune

grasses, thus forming sand dunes.
(Point out sand bars. Add that swimmers should obey the rules of life guards because Lake Michigan has frequent hazardous conditions. Describe rip currents, and what to do if you're caught in one.)

STOP 7 (Upper beach)

(Point out upper beach; mention plants that would live there if it weren't for people trampling them (Sea Rocket, Bugseed, Sand cress). Mention that many other animals -- Fowlers toads, raccoons, box turtles -- visit the beach at night. Birds visit by day, but can be scared away by people. Discuss habits, habitat and adaptations of herring gulls). (Walk over the foredune - stop at the lee side.)

STOP 8 (Foredune)

(Discuss Dune Succession)

1) Dune succession is the gradual change of plant communities growing on dunes. We will explore five of these communities. (Compare to changes in their home neighborhoods.) We have already seen the beach: this is the foredune community.
2) Marram grass is called a "pioneer" plant (discuss pioneers). It is usually the most common plant on forefunes. Its web-like roots stabilize the dunes. Well adapted

to desert-like conditions, marram grass roots extend into the ground to reach moisture. It's underground growth system (rhizomes) enables it to spread underground. It can push its leaves into the air if sand piles up around it, but if the sand erodes away, it will die. (Point out other types of dune grasses.) (Along the trail, point out examples of trail erosion.)

3) What word describes the ways marram thrives in difficult conditions? (adaptation)

4) Marram grass is a hardy plant, but walking on marram grass will destroy it, so please stay on the designated trail. Remember, this grass is extremely important in the process of dune building.

5) Look for digger wasp or burrowing spider holes. Discuss adaptations to living on hot, dry sand.

STOP 9 (near cottonwoods)

Activity:

"Eyes in your feet," soil sample #2

1) The cottonwood community is next. Cottonwoods can live in a variety of areas. Like marram grass, cottonwood trees are dune builders. New roots may start from a branch of a cottonwood tree if covered by an active dune. (illustrate using body and arms)

2) Briefly point out other plants and flowers (bearberry, goldenrod, dune cherry)

3) Question: What types of animals might one find in the cottonwood community? (Discuss Fowlers toads -- habit, habitat, adaptations).

STOP 10 (Jack Pine Forest)

Activity:

Soil Sample #3

1) The pine community follows the cottonwoods.

2) Here the ground has a richer topsoil than the previous communities (because of decay of previous generations). This community is also better shaded than the previous ones (providing better moisture retention and escape from the heat).

3) Could fire be useful to a forest? How? (Discuss fire ecology, especially relating to jack pines).

4) (Point out poison ivy and hop tree)

5) Discuss hog nose snake (habits, habitat, adaptations).

STOP 11 (Interdunal Pond)

(Discuss interdunal ponds):

1) Interdunal ponds are found in blowout areas. Dunes are sometimes eroded to point where the surface of the ground is at the water table. (Explain - the level of water underground). As it rains, the blowout fills with

water from the rain and from the ground, thus forming an interdunal pond.

2) (Discuss pond food webs -- the relationships among species found in and around ponds: mosquitoes, frogs and toads, tadpoles, insects, snakes, hawks, foxes, people).

3) (Discuss human role in food webs. Discuss adaptations.)

4) All things in nature are interrelated, directly or indirectly. There are controls and balances in food webs that if carelessly disrupted may eventually cause serious problems for humans and other species. (Discuss examples of possible disruptions.) Remind participants of how delicate the pond is, and to walk carefully around it. Tell them why we can't wade in the pond).

STOP 12 (Blowout I) (Impact is discussed here.) (Walk up on left side.)

1) (Discuss blowouts - a form of erosion) Question: What causes a blowout? (Anything which leads to the destruction of the vegetation on a dune: hiking off trail, fire, bicycles and motorcycles, animal trails, trees falling, lightning)

2) (Point out plants growing where people haven't walked: recall beach impact.)

STOP 13 (Cottonwood/erosion) (deleted in program)

(Point out branches sprouting from cottonwood roots; recall erosion problem, adaptation.)

STOP 14 (Panorama I)

(Stop and look back at the Lake and dunes. Review and locate the first four communities. Have group identify various signs of human impact (erosion, litter, industry, houses) Discuss ways to minimize impact. Be sure to discuss littering, staying on trails, not picking flowers, not bothering animals or visitors)

STOP 15 (Oak Forest)

Activity:

"Soil Study" Collect final sample

1) This is the community that succeeds the pines.
2) The oak forest generally has a richer top soil than the previous plant communities discussed (decay from previous communities).
3) Acorns need shade and moisture to sprout; they do not fare well in hot dry areas; (they grow in the shade of other plants, then take over). Suitable conditions are provided by each plant community for the "succeeding" one.
4) (Point out flowing sand on the right. Ask if they know what caused it. Leave unanswered until STOP 17.)

5) (Discuss raccoons -- habits, habitat, and adaptations.)

STOP 16 (Panorama II)

(Pause on the steps and let them look around. Ask them what they like and what they don't like. Ask them what this may look like in 5 years. Ask what they're willing to do to help protect it.)

STOP 17 (Blowout II)

(Look at blowout and slope. Ask why basswoods on top are dying -- erosion -- and what we can do about it -- don't climb dunes topped by forests. Connect erosion with sand noticed in STOP 15.

STOP 18 (Summary)

(Form a circle off the trail near STOP 18. Have them watch where they stand; stress that they are allowed to leave the trail with a Ranger because they've learned to respect the area. Review major points. Ask them what they learned, what they liked. Ask them what they will do when coming back. Questions, Thank you, and back to the bus.)

APPENDIX C: INSTRUMENTS

Name_____

Your Day with E.T.

A National Park Service Ranger, Rebecca Ranger, is leading your class on a field trip at a nature park. Suddenly, she stops at the beach edge. A strange thing is coming toward you and your class. You are excited because you think it's a creature from outer space! The thing comes nearer. Yes!! a creature from outer space has just landed in the park. It is E.T.! E.T. tells Rebecca that he wants to go on your field trip. Rebecca agrees and you're thrilled. But, E.T. doesn't know much about the nature park. E.T. needs you to help him. Fill in the correct word in the story below. HINT: Spelling words are listed at the end of the story. The number of blanks tells you the number of letters that the correct word has. You may need to use the same word two times and all the words that you need are listed with the spelling words.

"Urgh, urgh," says E.T. as he points his stick-like finger to the sand hills.

"-----," replies Rebecca.

Your class, Rebecca and E.T. next walk from the water to the upper beach area. Here, Rebecca Ranger points to the ground where --- ------ would grow if people didn't walk in the area. You keep walking. E.T. wonders why no plants are growing in the next community.

"People walk on these plants and they eventually ---. We worry because then the sand ----- away. One important plant that helps hold the dunes in place is ------ -----."

Figure C1. Knowledge Story

Name_____

"Urgh, urgh," exclaims E.T. as he points to a jumping animal which is among the cottonwood trees and wildflowers.

"That's a ------- ----," you tell E.T. "It is one of many ------- that lives here in the park."

You all continue along the path. Rebecca Ranger points out that many kinds of ------ grow in the different areas.

As you walk, E.T. points to another park visitor. The visitor is trying to catch a tadpole.

"We can't catch animals or take plants ---- because we need to leave them here for everyone," Rebecca explains to E.T. and your class.

E.T. nods.

"People also like to throw ----- at the animals," comments Rebecca Ranger.

You keep walking from the evergreens into a woodsy area, called the --- ------. By now you are tired and have decided that you want to go home and take E.T. with you.

You all walk up a long, steep hill and then down another hill to the bus. E.T. gets very excited when he sees the bus. You run over to the bus and E.T. follows right behind you. Your wish has come true today. E.T. is going to come home with you!

SPELLING WORDS

sand cherry
die(s)
dogwood tree(s)
hole(s)
bird(s)
oak forest(s)
box turtle(s)
lie(s)
home(s)
reptile(s)
sea rocket(s)
animal(s)
marram grass(es)
plant(s)
dune(s)
Fowler's toad(s)
goldenrod(s)
rock(s)
stick(s)
blow(s)

Figure C1 (Cont'd.)

BEHAVIOR

Please observe the children with the yellow nametags. Record the number on the child's nametag in the blank next to "Child." We appreciate your accuracy when you are recording their behaviors. Thank you very much.

1. At the designated site, circle whether the child is ON or OFF the trail. If possible, please suggest why the child was off the trail. (Fox example, trail too narrow, trying to get to head of line, looking for something off the trail, etc.)

			REASON
CHILD # ____	ON	OFF	_____
CHILD # ____	ON	OFF	_____
CHILD # ____	ON	OFF	_____
CHILD # ____	ON	OFF	_____
CHILD # ____	ON	OFF	_____
CHILD # ____	ON	OFF	_____

2. After the children are given something to eat, check "X" in the box that best describes what they do with the wrapper. Again, please fill in the number found on the child's nametag.

	Child___	Child___	Child___	Child ___	Child___	Child ___
Puts paper in pocket.						
Puts paper on trail.						
Puts paper in trashcan.						
Gives paper to another child.						
Other						

3. Circle if the child steps ON or OFF the main trail when s/he returns to the parking lot at the end of the program. Please also fill in the number found on the child's nametag.

child #____	ON	OFF	child #____	ON	OFF
child #____	ON	OFF	child #____	ON	OFF
child #____	ON	OFF	child #____	ON	OFF

Figure C2. Observation Checklist

Welcome to Indiana Dunes National Lakeshore. We are designing some new programs and are interested in finding out what specific content each of our programs cover. Please check in the blanks below if the specified information is discussed. The checklist is arranged by stops so you can keep track of where we are in the program. At some stops we discuss habits and habitats of various animals and plants. We will discuss habits in terms of what these animals eat and when they are active. Habitats are discussed when we talk about where the animals and plants live. An "Other" category is also provided for you to list additional information that might be covered at that particular stop. As we shall use this information in our planning, we appreciate your being as accurate as possible. Thank you very much for your time.

Figure C3. Content Checklist

STOP 1 - Sidewalk

_____Defines adaptation

Rules:

_____Stay on trails

_____Enjoy flowers and plants in their place

_____No littering

_____Watch animals from a distance

_____Do not disturb animals

_____Be courteous (ex. no talking while others are talking.)

_____Other_____

STOP 2 - North Area Parking Lot/Visitors' Center

ACTIVITY GROUP ONLY

_____Snake _____Habits _____Habitat

_____Turtle _____Habits _____Habitat

_____Other_____

STOP 3 - Beach House Road

_____Defines erosion

_____Points out erosion

_____Other_____

STOP 7 - Upper Beach

_____Defines Upper Beach

_____Searocket _____Habits _____Habitat

_____Bugseed _____Habits _____Habitat

_____Sandcress _____Habits _____Habitat

_____Fowler's toad and raccoon come here at night

_____Herring Gull _____Habits _____Habitat

_____Other_____

STOP 8 - Foredune

____Defines succession

____Defines foredune

____Marram grass ____Has runners ____Pioneer
____Habits ____Habitat

____Sand reed

____Blue stem

____Compares dune grasses with grasses at home

____Digger wasp ____Habits ____Habitat

____Burrowing Spider ____Habits ____Habitat

____Other _____

STOP 9 - Cottonwoods

____Defines cottonwood community

____Cottonwood ____Habits ____Habitat

____Bearberry ____Habits ____Habitat

____Goldenrod

____Fowler's toad ____Habits ____Habitat

____Other_____

STOP 10 - Jack pine Forest

____Defines jack pine community

____Jack pine ____Habits ____Habitat (includes fire)

____Poison Ivy ____Habits ____Habitat

____Hop tree ____Habits ____Habitat

____Hognose Snake ____Habits ____Habitat

____Other _____

STOP 11 - Interdunal Pond

____Defines interdunal pond

____Defines Blowout

____Defines watertable

____Other_____

STOP 12 - Blowout I

Blowout causes:

____Hiking off trail ____Bikes & Motor Bikes

____Trees falling ____Lightening

____Fire ____Animal trails

____Other_____

STOP 14 - Panarama I

Mention:

____Beach Community ____Cottonwood Community

____Foredune Community ____Jackpine Community

Human Impact:

____Litter ____Houses

____Erosion ____Industry

Ways to minimize impact:

____Stay on trails or flat open sand

____Pick up trash

____Don't bother animals

____Other_____

STOP 15 - Oak Forest

____Oak trees ____Habits (acorns, need richer soil) ____Habitat

____Raccoon ____Habits ____Habitat

____Other_____

STOP 17 - Blowout II

____Basswood dying

____Other_____

APPENDIX D: COVER LETTER TO DISTANT SCHOOLS

United States Department of the Interior
NATIONAL PARK SERVICE
INDIANA DUNES NATIONAL LAKESHORE
1100 N. MINERAL SPRINGS ROAD
PORTER, INDIANA 46304

IN REPLY REFER TO:

October 25, 1983

Dear Ms. Scott and Ms. Simons,

As part of our program development this fall at the Lakeshore, we wish to find out how familiar your students are with parks and Indiana Dunes National Lakeshore in particular. Michigan State University has helped us develop some stories, questions,and paper and pencil games in order for us to elicit this information. We would like your students to complete the sheets that we have sent. It will take about half of one hour. The sheets are divided into packets of 68. Please have each child fill out a green, gold and yellow sheet. The instructions are as self-explanatory as possible, yet past experience has shown that we need to clarify some points.

1. Have children put their names on all the sheets.

2. Clarify questions on the different sheets.

GREEN - Question 8 - choose one answer
Question 9 - answer "yes" or "no", then follow the arrow and answer the second part of the question
Question 10 - may have more than one answer

GOLD "Searching the Dunes"
last spelling word is "hognose snake"
answers may be singular or plural
3 down - the last block is marked out because of a print mistake

YELLOW "A Day at the Park"
have them choose the answer that best describes what they would do

We would like your students' input in developing our new programs and appreciate the time you take to have the children answer the questions. Please stress that the sheets are not tests and that their help will aid us in making our programs better meet everyone's needs.

United States Department of the Interior
NATIONAL PARK SERVICE
INDIANA DUNES NATIONAL LAKESHORE
1100 N. MINERAL SPRINGS ROAD
PORTER, INDIANA 46304

IN REPLY REFER TO:

When you arrive we would like you to have split each group of 65 into two groups (32 and 33 students each). We also want to make color tags (instead of name tags) for the students. Therefore, we need you to send us a class roster for both groups of 65 that shows the division of 32-33 students as soon as possible.

Please bring the completed sheets with you on November 2, 1983.

Thank you again very much for your time.

Sincerely,

Roland Hesselbart,
Environmental Education
Section, Indiana Dunes
National Lakeshore

Charlotte Young,
Michigan State University

APPENDIX E: RESPONSES FROM OPEN ENDED QUESTIONS

Table E1. Responses to the Open Ended Questions on the Attitude Questionnaires

Post-test Questionnaire:

Content of Response	Number of responses
Question #16: What do you think is the most important thing you learned today"	
not to make new trails	1
everything	1
how things live	1
beach	1
not kill animals	1
animals and beach	1
run down hill	2
woods and other plants	1
how some plants look	10
about poison ivy	1
dunes	1
soil	1
that you shouldn't litter	1
don't walk on plants' and grass	6
don't talk while ranger is talking	1
don't destroy nature	5
how to care for plants and animals	1
dunes, plants and animals	1
animals', plants' and insects' environments	1
plants and animals	8
about sand, trees and grass	1
animals	5
about dunes and plants	1
do not pick plants	2
leaves, plants, sand, rocks	3
nature	2
not to kill plants	1
sand	4
grass	1
snake	3

Post-test Questionnaire:

Question #17: What did you like best about your trip today?

the hills and the water	2
climbing up and down the mountains	7
walking and running on hills'dunes	24
water'lake	5
beach	3
learning	3
plants and trees	1
running and playing	1
trails	3
how lovely it was	1
looking at things I never saw before	1
dunes	8
learning about nature and West Beach	1
animal tracks	1
lake and waves	1
meeting nice people	1
talking	1
going to visitor center	1
everything I could do without hunting anything	1
all about hiking	1
sliding down hills	1

Responses to Open-ended Questions (Control and Experimental Group)

Post-test Questionnaire:

Question #10: What would you like to do if you came back to Indiana Dunes National Lakeshore?

slide'glide down the dunes	3
listen to birds	1
see animals	2
look at Lake Michigan	1
hike	4
look at snake	1

Question #10 (Continued)

run down hills without plants	2
look at things and talk about them	2
roll down hills	2
eat later	1
learn	1
swim	10
run, swim, play in sand	1
see inside of beach house	1
play/have fun	3
do same thing as today	5
clean up garbage	1
come back again	1
run on (dune) trails	9
ride my motorcycle	1
go down hills	27
climb on dunes (sandhills)	14

LITERATURE CITED

LITERATURE CITED

Anastasi, Anne. 1968. Psychological Testing. Toronto: Macmillan Co.

Anderson, S. B., and Ball, S. 1978. The Profession and Practice of Program Evaluation. San Francisco: Jossey-Bass Publications.

Babbie, E. 1973. Survey Research Methods. California: Wadsworth Publishing Co., Inc.

_____. 1983. The Practice of Social Research. 3rd edit. California: Wadsworth Publishing Co., Inc.

Bernard, N. T., Jr. 1977. "Earning Support for Interpretation." The Journal of Interpretation. 2(1):26-27.

Blahna, D. J. and Roggenbuck, J. W. 1979. "Planning Interpretation 'In Tune' With Visitor Expectations." In Proceedings, 1979 AIN Workshop, Bloomington, MN. pp. 5-7.

Bloom, B. S. and others. 1956. Taxonomy of Educational Objectives. Handbook I: The Cognitive Domain. New York: David McKay Co.

Bochner, S. 1979. "Designing Unobtrusive Field Experiments in Social Psychology." New Directions for Methodology of Behavioral Science.

Burgess, R. L., Clark, R. N. and Hendee, J. C. 1971. "An experimental analysis of Anti-littering procedures." Journal of Applied Behavior Analysis 4:71-75.

Burrus-Bammel, L. L. and Bammel, G. 1979. "Systematic Unobtrusive Measures." in Symposium on Evaluation Strategy: Assessing Outdoor Program Effectiveness. (compiled by Betty van der Simissen, HPER Series #12, College of Health, Physical Education and Recreation, Pennsylvania State University, St. College, PA. 1979.)

Campbell, D. T. and Fiske, D. W. 1959. "Convergent and Discriminant Validation by the Multi-trait Multi-method Matrix." Psychological Bulletin. 56(81-105).

Campbell, D. T. and Stanley, J. C. 1963. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally Co.

Cherem. G. J. 1975. "The Environmental Interpreter: New Frontiers," Paper presented at 14th Annual Meeting of the Association of Interpretive Naturalists, Natural Bridge, VA.

_____. 1978. "Evaluation of Interpretive Planning Efforts." Proceedings of the First Interpretation Central Training Institute. Interpertation Central, Ann Arbor, MI. pp. 59-62.

_____. 1983. in Fritschen "Interpretation for Management" (in press). Supplement to A Guide to Cultural and Environmental Interpretation in the U. S. Army Corps of Engineers. (Prospt. D. B. and Roggenbuck, J. W. 1981.)

Chubb, M. and Chubb, H. R. 1981. One Third of Our Time? New York: John Wiley & Sons, Inc.

Clark, R. N. 1977. "Alternative Strategies for Studying River Recreationists." in Proceedings 1977, River Recreation Management and Research Symposium. Minneapolis, MN. pp. 91-100.

Clark, R. N., Hendee, J. C., and Burgess, R. L. 1972. "Development of Anti-littering behavior in a forest campground," Journal of Applied Behavior Analysis 5:1-5.

"Cooperative Extension Service,". Journal of Extension. 1979. (see Quick and Davis citing)

Corps of Engineers. 1983. Research conducted by Roggenbuck, Cherem and Freed for U. S. Army Corps of Engineers.

Crandell, R. and Slivken, K. 1980. "Leisure Attitudes and their Measurement" in Iso-Ahola, S. (edit.) Social Psychological Perspectives on Leisure and Recreation. Illinois: C. C. Thomas Publisher.

Crano, William and Brewer, Marilynn. 1973. Principles of Research in Social Psychology. New York: McGraw-Hill.

Crompton, J. L. and Sellar, C. (no date) "Do Outdoor Education Experiences Contribute to Positive Development in the Affective Domain: A Review of the Literature." report, Department of Recreation and Parks, Texas A & M University, Texas.

Dartington Amenity Research Trust & Dept. of Psychology, Univ. of Surrey. 1978. Interpretation at Visitor Centres. Cheltamham, ENG: Crown Copyright.

Dawson, M. S. and Roggenbuck, J. W. 1979. "Evaluation of Historical Interpretation in National Park Service Areas." In Proceedings, 1979 AIN Workshop, Bloomington, MN., pp. 15-18.

Dick, R. E., Myklestad, E. and Wagar, J. A. 1975. Audience attention as a basis for evaluating interpretive presentations. U.S.D.A. Forest Service Research Paper PNW-198. Pacific Northwest Forest & Range Experiment Station, Portland, ORE.

Elliot, D. L. 1979. "Training Interpreters and Environmental Educators: Experiences Resulting in Attitude Change." In Proceedings, 1979 AIN Workshop, Bloomington, MN. pp. 18-21.

Environmental Education Program Evaluation. 1981. ERIC, ERIC'SMEAC Fact Sheet #1, Ohio.

Fishbein, M. and Ajzen, I. 1975. Belief, Attitude, Intention and Behavior. Massachusetts: Addison-Wesley Publishing Co.

Freed, M. 1983. in Fritschen "Interpretation for Management." (in press), Supplement to A Guide to Cultural and Environmental Interpretation in the U.S. Army Corps of Engineers. (Propst. D. B. & Roggenbuck, J. W. 1981.)

Fritschen, J. M. 1980. An Evaluation of Two Methods of Relating Environmental Interpretation to Urban Residents of Detroit, MI. unpublished M.S. Thesis, Department of Parks and Recreation Resources, Michigan State University, E. Lansing, MI.

_____. 1983. "Interpretation for Management" (in press) Supplement to A Guide to Cultural and Environmental Interpretation in the U.S. Army Corps of Engineers. (Propst, D. B. & Roggenbuck, J. W. 1981.)

Garrett, H. E. 1960. Statistics in Psychology and Education. New York: Longmans, Green & Co.

Gramann, J. H. and Field, D. R. 1977. Interpretive Activity Inventory: Analysis of the 1976 Data from Mt. Rainier National Park and Perry's Victory and International Peace Memorial, National Park Service, Cooperative Park Studies Unit, College of Forest Resources, Univ. of Washington, Seattle.

Gustke, L. D. and Hodgson, R. W. 1980. "The Effect of an Environmental Discontinuity on the Rate of Travel along an Interpretive Trail." Environment and Behavior 12(1):53-63.

Haire, M. 1950. "Projective techniques in marketing research." Journal of Marketing. 14:649-656.

Ham, S. H. and Shew, R. L. 1979. "A Comparison of Visitors' and Interpreters' Assessments of Conducted Interpretive Activities." J. of Interpretation. 4(2):39-44.

Hammitt, W. 1981. "Attention, Familarity and Effective Interpretation." paper presented at AIN 1981, National Conference - Estes Park, Colorado.

Hodgson, R. W. 1977. "Building Evaluation into Interpretation" Presented at Association of Interpretive Naturalists 1977 National Workshop, College Station, Texas.

Hodgson, R. and Fritchen, J. 1982. "Evaluating Interpretation" Supplement to A Guide to Cultural and Environmental Interpretation in the U.S. Army Corps of Engineers. (Propst and Roggenbuck, 1981).

Hoffman, B. 1967. "Towards less Emphasis on Multiple Choice Tests" in Educational and Psychological Measurement Payne, D. A. and McMorris,(edits.)Massachusetts: Blaisdell Publishing Co.

Houston, T. R. Jr. 1972. "The Behavioral Sciences Impact-Effectiveness Model." (in Evaluating Social Programs. Rossi, P. H. and Williams, W. (edits.) New York: Seminar Press.)

Iozzi, L. A. and Cheu, J. M. Y. 1979. "Assessing Outdoor/ Environmental Programs. The Cognitive-Developmental Approach." in Symposium on Evaluation Strategy: Assessing Outdoor Program Effectiveness compiled by Betty van der Smissen P.S.U. HPER, Series #2, Coll. HPER, 1979.

Kazdin. 1979. "Direct Observations as Unobtrusive Measures in Treatment Evaluation." New Directions for Methodology of Behavioral Science.

Kidder, Louise, H. (edit.). 1981. Selltiz, Writghtsman and Cook's Research Methods in Social Relations, 4th edit. New York, Holt, Rinehart & Winston.

Knudson, D. and Morfoot, C. 1979. "Cost Effectiveness of Interpretation in Indiana Recreation Areas." in Proceedings AIN Annual Meeting, Bloomington, MN.

Kuehner, R. 1972. "Visitor Behavior at the Lake Tahoe Visitor Center: A Preliminary Report of an Observation and Interview Study of the Activities of Vacationers on a Forest Service, Self-Guided Nature Trail,: Forest Recreation Research Project, Berkeley, CA.: Pacific Southwest Forest and Range Experiment Station.

LaHart, D. E. and Bailey, J. S. 1975. "Reducing Children's Littering on a Nature Trail. "Journal of Environmental Education. 7(1):37-45.

Lamb, C. W. and Crompton, J. R. 1980. "Qualitative Measures of Program Success." in Recreation and Leisure: Issues in An Era of Change. (Goodale, T. L. and Witt, P. A. edits.) State College, PA.: Venture Publishing.

Lange, R. R. 1980. Environmental Education Needs Assessment and Evaluation Manual, Vol. 1 & 2. prepared for Colorado Department of Education.

Lansing, J. and Heyns, R. 1959. "Need affiliation and frequency of four types of communication." Journal of Abnormal and Social Psychology 58:365-372.

Lauffer, A. 1978. Social Planning at the Community Level. New Jersey: Prentice-Hall, Inc.

Lee, T. R. and Uzzell, D. L. 1980. The Educational Effectiveness of the Farm Open Day at Scotland, ENGLAND, Country Side Commission.

Lemke, E. and Wiersma, W. 1976. Principles of Psychological Measurement. Chicago: Rand McNally Publishing Co.

Lesser, G. S. 1974. Children and Television. N.Y.: Vintage Books.

Lime, D. W. 1979. "Visitor Observation: A Tool for Appraising Interpretive Activities," In Proceedings, 1979 AIN Workshop, Bloomington, MN. pp. 49-54.

Lipman, D. and Hodgson, R. 1978. "The Influence of Interpersonal Interpretation on the effectiveness of self-guided cave tours." Journal of Environmental Education 10(1):8.

Loether, H. J. and McTavish, D. G. 1974. Descriptive Statistics for Sociology. Boston: Allyn & Bacon, Inc.

Magnusson, D. 1967. Test Theory Reading, Mass.: Addison-Wesley Publishing Co.

Mahaffey, B. 1969. "Relative Effectiveness and Visitor Preference of Three Audio-Visual Media for Interpretation of a Historic Area." Departmental Information Report #1, Department of Recreation and Parks, Texas A&M University, College Station, TX.

McDonough, M. H. 1981. "Interpretation Methods and Devices." Department of Parks and Recreation Resources University Course Lecture, Michigan State University, E. Lansing, MI.

McGaw, D. and Watson, G. 1976. Political and Social Inquiry N.Y.: John Wiley & Sons.

Meis, S. and Davis, G. 1981. The Evaluation of an Unwanted Interpretive Exhibit." Paper presented at the Sixth Annual Meeting of the Canadian Association of Applied Social Research, Learned Societies Conference. Halifax, Nova Scotia, Canada.

Moeller, G. H., et al. 1980. "The Informal Interview as a Technique for Recreation Research." Journal of Leisure Research. 12(2):174-182.

Monroe, M. 1983. Evaluation studies conducted at Dahlem Environmental Education Center, Jackson, MI.

Morfoot, C. F. and Blake, B. F. 1979. "Interpretive Effectiveness Improving Techniques of Evaluation." In Proceedings, 1979 AIN Workshop, Bloomington, MN. pp. 57-59.

Moses, Epstein and Wiseman, Inc. 1977. "Assessing the Impact of Interpretive Programs." unpublished report #PX-001-07-0702, prepared for Division of Interpretation and Visitor Services, National Park Service. Washington, D.C.

National Park Service Interpretive Effectiveness Research Project. 1983. Funded by Eastern Parks and Monuments Association.

Nie, N. H., Hull, C. H., Jenking, J. G., Stienbrenner, K. and Bent, D. H. 1975. Statistical Package for the Social Sciences. 2nd edit. New York: McGraw-Hill Book Co.

Nielsen, C. 1980. "An Evaluation of Public Interpretation Programs as an aid to Resource Management." Grand Teton National Park.

Peart, B. 1978. "Definition of Interpretation," Interpretation Canada 5:2(3-6), Aylmer, Quebec, Canada.

_____. 1979. "Evaluation Studies of Two Canadian Wildlife Interpretive Centres." in Proceedings, AIN Workshop, Bloomington, MN. pp. 62-63.

Peyton, B. 1982. "Outdoor Environmental Education," Department of Fisheries and Wildlife. University Course Lecture, Michigan State University, E. Lansing, Michigan.

Poister, T. H. 1978. Public Program Analysis: Applied Research Methods. Baltimore: University Park Press.

Powers, R., Osborne, J. and Anderson, E. G. 1973. "Positive Reinforcement of litter removal in the natural environment." Journal of Applied Behavior Analysis. 6:579-586.

Propst, D. B. and Roggenbuck, J. W. 1981. A Guide to Cultural and Environmental Interpretation in the U.S. Army Corps of Engineers. Instruction Report R-81-1 Vicksburg, Mississippi. Waterways Experiment Station.

Putney, A. D. and Wagar, J. A. 1973a. "Objectives and Evaluation in Interpretive Planning," Journal of Environmental Education. 5(1):43-44.

Putney, A. D. and Wagar, J. A. 1973b. "Cost per visitor contact: a step toward cost effectiveness in interpretive methods." Pacific Northwest Forest and Range Experimental Station Technical Report.

Quick, S. and Davis, K. 1979. A Self-Reporting Evaluation Tool,: Journal of Extension IVII: 20-25.

Risk and English. 1983. "An Observational Strategy for Evaluating Environmental Interpretation." unpublished report Department of Recreation and Parks, Pennsylvania State University, State College, PA.

Robinson, J. P. and Shaver, P. R. 1973. Measures of Social Psychological Attitudes. (revised edit) Ann Arbor: Survey Research Center; Institute of Social Research.

Roggenbuck, J. W. 1979. "The Field Experiment: A Suggested Method for Interpretive Evaluation," Journal of Interpretation. 4(1):9-11.

Roggenbuck, D. 1983. in Fritschen "Interpretation for Management." (in press). Supplement to A Guide to Cultural and Environmental Interpretation in the U.S. Army Corps of Engineers. Propst, D. B. and Roggenbuck, J. W. 1981.

Rossi, P. H. 1972. "Testing for Success and Failure in Social Action." (in Evaluating Social Programs. Rossi, P. H., and Williams, W., (edits.) New York: Seminar Press.)

St. Clair, S. L. Jr. 1972. "Relative Effectiveness of Two Personal Interpretive Methods at a Community Nature Center," unpublished M.S. thesis, Department of Parks and Recreation Resources, Michigan State University, East Lansing, Michigan.

Schafer, R. B. and Tait, J. L. 1981. A Guide to Understanding Attitudes and Attitude Change. North Central Regional Extension Publication, 138, Ohio State University, Extension Sociology Committee.

Sharpe, G. W. 1982. Interpreting the Environment. 2nd edit. New York: John Wiley & Sons.

Sharpe, G. W. and Gensler, G. L. 1977. "Interpretation as a management tool." paper presented at Assn. of Canadian Interpreters, Banff, Alberta.

Shaw, M. E. and Wright, J. M. 1967. Scales for the measurement of attitudes. New York: McGraw-Hill.

Shiner, J. W. and Shafer, E. L. Jr. 1975. "How Long do People Look at and Listen to Forest-Oriented Exhibits?" U.S.D.A. Forest Service Research Paper N-E 325. Northeastern Forest Experiment Station, Broomall, PA.

Sommer, R. and Sommer, B. B. 1980. Practical Guide to Behavioral Research. N. Y.: Oxford Univ. Press.

State of the Art on Methodology for Studying Environmental Perceptions, Beliefs, Attitudes and Values and State of the Art on Utilizing Perception, Attitude and Opinion Research. A synopsis of a conference held at Cornell Univ., by Rockefeller Foundation and Dept. of Rural Sociology, College Ag. and Life Sciences, Ithaca, N.Y.

Suchman, E. A. 1967. Evaluative Research. New York: Russell Sage Foundation.

Tai, Doris B. 1981. "An Evaluation of the Use and Effectiveness of Two Types of Interpretive Trail Media in Yellowstone National Park. unpublished Masters thesis, Univ. of Idaho, Moscow, Idaho.

Tilden, F. 1977. Interpreting our Heritage. N. C.: The University of North Carolina Press.

Theobald, W. F. 1979. Evaluation of Recreation and Park Programs. New York: John Wiley & Sons.

Wagar, J. A. 1972. "The Recording Quizboard: A Device for Evaluating Interpretive Services." U.S.D.A. Forest Service Research Paper PNW-139. Pacific Northwest Forest and Range Station, Portland, ORE.

_____. 1976. "Evaluating the Effectiveness of Interpretation", Journal of Interpretation. 1(1):8pp.

_____. 1980. "Effectiveness in Interpretation: Let OPEC Be Our Guide." In Resource preservation through interpretation, Proceedings California State Park Rangers Association, California Workshop-Convention, Santa Cruz, CA. pp. 42-48.

_____. 1982. Research in Interpretation. In Sharpe, G. W. Interpreting the Environment. New York: John Wiley & Sons.

Wagar, J. A., Lovelady, G. W. and Falkin, H. 1976. "Evaluation Techniques for Interpretation: Study Results from an Exhibition on Energy," U.S.D.A. Forest Service Research Paper PNW-211 Pacific Northwest Forest and Range Experiment Station, Portland, ORE.

Washburne, R. F. and Wagar, J. A. 1972. "Evaluating Visitor Response to Exhibit Content." Curator. 15(3):248-254.

Webb, E. J., Campbell, D. T., Schwartz, R. D. and Sechrest, L. 1970. Unobtrusive Measures: Non-reactive Research in the Social Sciences. Chicago: Rand McNally.