THE EFFECTIVENESS OF TRAINING RIVER GUIDES AS AN ALTERNATIVE INTERPRETIVE APPROACH IN THE NEW RIVER GORGE

by

Clifton T. Bobinski

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE
in
Forestry

APPROVED:

J.W. Roggenbuck, Chairman

D.E. Cockrell

J.D. Wellman

R.E. Adams-Dept. Head

June, 1985

Blacksburg, Virginia

THE EFFECTIVENESS OF TRAINING RIVER GUIDES AS AN ALTERNATIVE INTERPRETIVE APPROACH IN THE NEW RIVER GORGE

by
Clifton T. Bobinski
Committee Chariman: Joseph W. Roggenbuck
Forestry

(ABSTRACT)

The effectiveness of a guide training program was evaluated as a means of providing interpretive services to commercial boaters at the New River Gorge National River. Commercial river guides attended a National Park Service sponsored training program which provided accurate information and education concerning natural and cultural history of the New River, the national significance of the New River Gorge National River, the history and purpose of the National Park Service, and information and services available at the visitor centers. The emphasis of the training session was to increase the river guides' knowledge base and to encourage their interpretive presentation of this information to their customers.

Customers of a commercial outfitter were administered a questionnaire before and after the guide training program. Significant differences in the amount of interpretation guides presented on the river, the amount of knowledge customers acquired during the trip, and the customers' overall trip rating were noted by empirical testing. Increases in the means of all three outcome variables occurred following guide training. Customers' intentions to visit a New River Gorge Visitor Center did not significantly change following guide training. The validity and reliability of the instrument is discussed as well as potential biases and constraints of the study. Implications for management and further research are also discussed.

ACKNOWLEDGEMENTS

This project was a test of personal will and fortitude that was made possible by the assistance of several individuals. Many friends, but perhaps foremost Tim Easley, Dave Turner, and Tom Waldron, were always supportive and caring. Jerry, Randy, Charlotte, Annette, and Tim made life much more fun and enjoyable when the situation seemed otherwise. The mysteries of the mainframe and statistics would still be unraveling if it weren't for Jim, Kurt and David.

A special thanks is due to all the folks at Mountain River Tours and the National Park Service who either participated in the study or helped with the design and administration of the guide training program. Gene Cox and Neil Dejong in particular deserve to be mentioned for contributing so much time, energy, and assistance.

I would like to thank Joe, Doug, and Dave, my committee, for all the good suggestions, feedback, and advice. While times were often more than challenging and quite trying, I did learn a great deal for which I am grateful.

Though not specifically mentioned, many other friends and family members contributed immensely to this effort. To them I will always be thankful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Interpretive programs are a traditional and often legally mandated service provided for visitors to the national parks. Typically these services have been provided by the National Park Service. However, in recent years, as the role of park concessionaires and private enterprise in providing visitor services has increased (Reagan, 1985), so has the involvement of the private sector in providing interpretive services. Today, with increased budget and manpower constraints and the trend toward privatization of government services, there is a great need for the study of both the feasibility and effectiveness of interpretive programs provided by the concessionaires and/or private enterprises operating on federal lands.

Commercial whitewater rafting is a very popular recreation enterprise now operating on numerous federally managed rivers. Whitewater rafting at New River Gorge National River in West Virginia, for example, attracted over 77,500 people during the 1983 season. The responsibilities of the National Park Service, the government agency charged with the administration and management of this National River, as described by Public Law 95-625, are:

-conserving the outstanding natural, scenic, and historic values and objects in and around the New River Gorge;

-interpreting the resources for the visiting public;

-preserving the free-flowing river for the benefit and enjoyment of present and future generations.

The responsibility for providing interpretive services to the whitewater visitor then lies with the National Park Service. However, due to the large number of visitors, the limited manpower, and the remote and dispersed nature of the boating activity, the Park Service cannot depend solely on traditional interpretive means to meet this responsibility. While the National Park Service recognizes the importance of interpretation, on-site river interpretation has been left almost entirely to the commercial outfitters and their guides.

This National Park Service management strategy appears to be reasonable, for commercial river guides have historically been expected to provide their customers with information and education concerning the river and the river recreation experience. Many commercial outfitters consider the communication ability of the guide to be a critical employment criterion (Mountain River Tours, 1984). Bange (1984) found guides on the New River to be very influential both as sources of information and in the development of norms and beliefs in visitors concerning appropriate use and use levels in the Gorge. Cockrell et al. (1984) found New River guides to have higher social influence over their customers on the river trip than family, other party members, and other referents. Bange (1984) found that guides can be effective communicators, that the extent of their influence as well as other desirable characteristics and behaviors are notable, and concluded that the enhancement of river experiences and management through the guide's

abilities should be considered for potential use by both Park Service managers and outfitters.

Essentially, some guides are already attempting on-site river interpretation but without formal training and support in the development of accurate information from the National Park Service. Perhaps for this reason, river guides may not be as effective as they could be if they had accurate information.

A recent study indicates that most commercial boaters on the New River never come in contact with the National Park Service (Roggenbuck and Bange, 1983), and interpretation of the river resources by the National Park Service is not generally an integral part of their trip. Most boaters get information about the New River from friends and acquaintances, and second most frequently, from the river outfitters. Even with two attractive and easily accessible visitor centers, the National Park Service is almost never a source of information for boaters, yet these same visitors expressed a desire for receiving more information and interpretation in the Gorge (Roggenbuck and Bange, 1983).

PROBLEM STATEMENT

A cooperative effort involving the National Park Service and the commercial whitewater outfitters appears to be the most promising means of providing interpretive services to New River visitors. A cost effective strategy may be for the National Park Service to continue to depend on the river outfitters for interpretive services, but to work with and improve the information base of the commercial river guides through a guide training program. The training program would provide guides with accurate information and encourage them to pass it along to their customers. Satisfaction of boaters' desires for more information as well as the legal mandate for interpretation might then be accomplished.

The Park Service has offered such a staff training program on a voluntary basis to outfitters in the past, but no evaluation has been done. This research addressed the problem of evaluating the effectiveness of a guide training program as a means of providing interpretive services to commercial boaters at the New River Gorge National River.

OBJECTIVES

The National Park Service is currently developing a River Management Plan to guide overall use and prevent a significant loss in the quality of a national park experience on the river. The River Management Plan will also recommend strategies to meet the legal mandate prescribing interpretive services for park visitors. This study evaluated one potential strategy.

The goal of this study was to determine whether river guides could increase knowledge and enjoyment levels of their guests (the park visitors) through their interpretive efforts on the river. The specific objectives were:

1. To determine if following a National Park Service sponsored training program for the guides, the amount of interpretation presented on trips increased.
2. To determine whether visitor knowledge of the National Park Service, the New River Gorge National River, the significance of the New River Gorge and the natural and cultural history along the river increased on trips following guide training.
3. To determine whether visitor intentions to visit a New River Gorge visitor center increased on trips following guide training.
4. To determine whether interpretation increased visitors' enjoyment of their trip.
5. To determine whether the effectiveness of the National Park Service sponsored interpretive training program and the

effectiveness of guide interpretation was affected by the experience level of the guide.

This study stressed development of the guides' message content and not the message delivery. It was assumed that each guide had a personal communication style, and it was not the intention of this study to manipulate this aspect of communication. Rather, informational and educational training was intended to increase the guides' knowledge, the amount of interpretation offered, and, in turn, the park visitor's knowledge.

LITERATURE REVIEW

Information and interpretive services are not only desired by the commercial boaters at the New River Gorge but mandated by law as a responsibility of the National Park Service (PL 95- 625). This review of literature is an attempt to answer the following basic questions:

1. What is interpretation? What is its purpose? What are its benefits? Can one expect interpretation to increase enjoyment, increase knowledge, and change behavioral intentions?
2. Can river guides provide effective interpretive services?
3. How does one evaluate the effectiveness of the interpretive services or training?

Several authors propose definitions of interpretation which cite various purposes and attainable benefits. Interpretation is defined by Edwards (1976) as attractive communication offering concise information given in the presence of the topic with the goal of revealing significance. Wagar (1978) described interpretation as a vital activity of utmost importance to society in that it provides increased recreational benefits from a resource in a nonconsumptive way. By helping recreationists enjoy and understand the areas they visit, interpretation of natural and cultural history can add substantially to the quality of visitor experiences and therefore to the stream of benefits produced by such areas (Putney and Wagar 1973). These benefits might consist of new insights otherwise missed or simply enjoyable experiences to the

non-captive visitor who is in a leisure frame of mind (Cherem 1977). Indeed, interpreters talk of informing and educating people, and of enhancing or enriching their experiences, but entertainment, fun, and pleasure are benefits that might also be expected (Wagar 1978). Deane (1979) reported very high enjoyment levels following guided interpretive walks, one of three alternative nature interpretation techniques examined by the study. Warder and Raile (1984) found that interpretive methods using personal services (versus non-personal services) were the method most enjoyed by park visitors. It thus seems reasonable to expect interpretation, and especially interpretation involving personal services, to increase visitors' enjoyment.

Interpretation is also defined as an educational activity which aims to reveal meanings and relationships through the use of original objects, by firsthand experience, and by illustrative media rather than simply to communicate factual information (Tilden 1967). In addition to providing a purely recreational experience, interpretation has the potential to provide visitors with information about the resource and to aid visitors in understanding resource management policies (Nielson 1981). Nielson (1981) evaluated two interpretive methods and found that both a slide-show treatment and an interpreter-guided treatment resulted in higher mean knowledge scores on a fire ecology scale. Warder and Raile (1984) reported a significant increase in knowledge of pertinent history at the Fort Laramie National Historic Site following use of several interpretive methods, such as personal services, attended services, and non-personal services. Personal services consisted of talks given by uniformed personnel and attended services consisted of talks given by historically

dressed personnel. Significant differences in the effectiveness of these different methods were reported and it was found that personal services resulted in a higher knowledge level than did non-personal services (Warder and Raile 1984). Fazio (1979) reported tests of the effectiveness of interpretation on increasing "wilderness knowledge" levels and found wilderness rangers to be very effective at increasing knowledge. Gaining knowledge from an interpretive program is thus a reasonable expectation and benefit of interpretation (Nielson, 1981).

Wagar (1976) suggests that interpretation may provide a hierarchy of benefits starting with audience attention to the presentation and progressing to retention of information, changed attitudes, persistent behavioral change, and appreciation and enjoyment. Authors concerned with attitude theory (Fishbein and Azjen, 1975; Kiely-Brocato, 1980) suggest a positive relationship among these benefits. It seems increased knowledge can foster attitude change and in turn behavior change. Nielson (1981) found that interpretive programs increased knowledge of fire ecology and increased support for natural fire management as well, evidence tending to support current attitude theory. Changes in campers' behavior was noted by Roggenbuck and Berrier (1982) who found both brochures and personal contact to be effective in dispersing wilderness campers from a heavily used meadow. Lime and Lucas (1977) have similarly suggested that simply providing wilderness visitors with good information increases their knowledge of available alternatives and their ability to make informed decisions. Lime and Lucas actually noted these changes in visitors to the Boundary Waters Canoe Area.

Actual changes in visitor behavior due to interpretive efforts have therefore been noted in the literature. Interpretation has also been found to be viable and often indispensable in accomplishing management goals concerning site preservation, visitor protection, and depreciative behavior and vandalism (Sharpe and Gensler, 1978). Oliver et al. (1985) reported an 80% decrease in depreciative acts in campsites following distribution of a brochure to campers through a personal contact.

Reasonably then, if a person's behavior is highly correlated with his intention to perform that behavior (Fishbein and Azjen, 1975) then interpretation which has changed some behavior must have changed some behavioral intentions as well. Intentions to perform a given behavior are determined by "attitudinal" and "normative" factors (Fishbein and Azjen, 1975). The fact that interpretation has been shown to change attitudes (Nielson, 1981) and that New River guides have been shown to exert a high degree of normative influence (Bange, 1984) suggests that interpretation may be used to change behavior and behavioral intentions. However, behavior itself as well as behavioral intentions are very complex (Young and McDonough, 1984), and therefore may prove difficult to change by simply exposing individuals to an interpretive program. Still, the expectation is that an interpretive message describing the benefits and services available at a New River Gorge Visitor Center will increase visitors' knowledge of these services and in turn their intention to visit the facility.

The tacit notion is that guides will be "interpreting" the New River for their customers. While Bange (1984) suggests that New River Guides are influential communicators, there is still the issue of whether guides are providing interpretation and whether training would have any consequences.

Support is found in Cherem's (1977) discussion of types of interpreters. He states:

> "The time has come for all individuals performing the interpretive function (like docents for museums and zoos or guides for commercial enterprises) to be called "interpreters"...while some intepreters today are highly qualified, many are not. To prepare interpreters...we must provide them with the best training possible. This training can be obtained through professional, rigorous college curricula, through short courses, or through in-service training in interpretation."

The underlying premise of the study reported here was that guides are "performing the interpretive function" of which Cherem speaks. They do this when they seek to offer accurate on-site information concerning natural and cultural history in order to increase their customers' understanding and enjoyment of the New River. The expectation is that a short training session offered by the National Park Service will better prepare and qualify the guides and improve their effectiveness.

For example, Oliver et al. (1985) suggest that a training program provided by project managers at a Corps of Engineers reservoir likely increased the effectiveness of the study's visitor-contact person. The visitor-contact person was provided training bout the agency's philosophy, area resources, facilities, and programs. This person distributed brochures while contacting visitors and proved highly effective in reducing depreciative acts among those campers (Oliver et al. 1985).

Research further suggests that the medium of interpretation the guides use, namely personal contact, is a most effective medium. In interpreting river resources, Harrison (1977) has identified the personal contact as by far the most flexible and effective method. Deane (1979) examined three alternative nature interpretation techniques and found the greatest

increase in audience knowledge following the guided walk. Other professional interpreters contend that personal and attended programs are preferable to other type of programs (Sharpe, 1982), and this is supported by empirical testing (Roggenbuck et al., 1982; Fazio, 1984; Warder and Raile, 1984).

Several authors suggest that the effectiveness of an interpreter may be substantially affected by his/her personal characteristics (Risk, 1982; Sharpe, 1982; Propst and Roggenbuck, 1981). Personality variables like poise, credibility, enthusiasm, self-confidence, and articulateness are some of the personal attributes of the interpreter which account for variations in success. According to Sharpe (1976) an interperter's articulateness can increase with practice and is molded by experience. Likewise, poise is a composite of several traits that includes maturity, confidence, and warmth. It is also a trait that will grow with experience and age (Sharpe, 1976). Roggenbuck and Berrier (1982) suggest that wilderness ranger effectiveness in dispersing campers may vary by the personal characteristics of sex, personality, and experience. Bange (1984) noted that customers who perceived guides as skilled and trustworthy also viewed them as more credible in the information they gave and more influential. It is therefore suggested that as guides become more experienced, they also will become more skilled, more articulate, more poised, more credible, and thereby more effective as interpreters.

Training is suggested as one means to increase interpreter effectiveness, particularly for inexperienced or untrained personnel (Hodgson, 1984; Oliver et al, 1985). This suggests that inexperienced guides may get more out of a guide training session than experienced guides. Intuitively one would expect inexperienced guides to also be

more inexperienced with the content of a guide training session than experienced guides. Further, inexperienced guides probably have more room for improvement from such a program, while experienced guides may exhibit a ceiling effect. An analogous situation was reported by Warder and Raile (1984) where visitors to a historic site who had made previous visits exhibited significantly higher knowledge scores than did first-time visitors. However, first time visitors showed significantly higher score increases following the interpretive program than did repeat visitors. The guide training program serves as an interpretive treatment which will probably have the greatest impact on those least experienced with its content.

Many authors have cited the need for the evaluation of interpretive programs to determine their effectiveness (Wagar, 1976; Roggenbuck, 1979; Nielson, 1981). Manuals evaluating recreation and park programs, environmental education, and cultural and environmental interpretation have been written to assist in the accomplishment of this task (Theobald, 1979; Lange, 1980; Propst and Roggenbuck, 1981; US Corps of Engineers, 1984). Relatively few evaluations of interpretive programs, however, have been published (Young, 1984), and literature addressing on-site river interpretation is nearly non existent (Harrison, 1977). The studies that have been done offer insight into the process of evaluation.

Researchers now define interpretive effectiveness from an audience perspective (Wagar, 1976) by setting program objectives which specify the intended impact on the audience (Young, 1984). Program effectiveness then refers to the extent to which the objectives of the program have been achieved (Theobald, 1979). Propst and Roggenbuck (1981) stressed that

program objectives are necessary for evaluation, and the best measure of program effectiveness is the audience response.

Wagar (1976) identified a framework for evaluating audience response to interpretation, and measures of enjoyment and retention of information were two components. Many other studies have operationally defined effectiveness as knowledge recall/retention (Wagar, 1972; Fritschen, 1980; Nielson, 1981; Dick, et al., 1975). Interpretive effectiveness can be defined and measured in a variety of ways, such as increases in enjoyment or knowledge, as well as behavior change, but more than one of these measures is advisable for thorough evaluation (Propst and Roggenbuck 1981).

Various evaluation techniques have been compared for usefulness and applicability (Wagar, 1976; Propst and Roggenbuck, 1981). Questionnaires can provide a large amount of valid and reliable data about the audience and can assess knowledge, opinion, attitude, and behavioral intentions (Peart, 1979; Fritschen, 1980; Nielson, 1981). While the questionnaire seems most appropriate for this study, Wagar (1976) warns against burdening the visitor. A short questionnaire appears best, and Roggenbuck (1979) suggests the use of a field experiment technique. Hodgson and Fritschen (1984) propose the use of experimental and quasi-experimental designs so that evaluation can measure visitor response while controlling for other extraneous factors.

HYPOTHESES

The rationale and basis for the hypotheses tested by this study was provided by the preceding review of literature and empirical research. This review permitted the following hypotheses to be stated in non-null form.

H-1: Following guide training, the amount of interpretation presented on the river will increase (Cherem, 1977; Oliver et al., 1985).

H-2: Following guide training, visitors' knowledge of the National Park Service, the national significance of the New River Gorge National River, and the natural and cultural history of the river will increase. (Fazio, 1979; Nielson, 1981; Warder and Raile, 1984).

H-3: Following guide training, visitor ratings of the quality of the trip will increase. (Wagar, 1976, 1978; Cherem, 1977).

H-4: Following guide training, visitors will show increased intentions to visit a New River Gorge visitor center (Wagar, 1976; Fishbein and Azjen, 1975; Nielson, 1981; Bange, 1984; Roggenbuck and Berrier, 1982; Oliver et al., 1985; Lime and Lucas, 1977).

H-5: Posttest knowledge scores of customers who had an experienced guide will be higher than posttest scores of customers who had an inexperienced guide, both before and after training (Sharpe, 1976;, 1982; Propst and Roggenbuck, 1981; Bange, 1984; Hodgson, 1984).

Guide training is hypothesized to have the greatest effect on the inexperienced guides. Therefore, prior to the training program, it is expected that the gain scores of visitors associated with experienced guides will be greater than those of visitors associated with inexperienced guides. The training program is expected to remove this inequality and, following training, result in no differences among mean gain scores of experienced and inexperienced guides. This rationale leads to the following hypotheses.

H-6a: Before training, pretest-posttest gain scores associated with experienced guides will be significantly different and greater than gain scores associated with inexperience guides (Hodgson, 1984; Oliver et al., 1985; Warder and Raile, 1984).

H-6b: Following training, pretest-posttest gain scores associated with experienced guides will not be significantly different from gain scores associated with inexperienced guides (Hodgson, 1984; Oliver et al., 1985; Warder and Raile, 1984).

METHODOLOGY

Study Area

The New River Gorge National River was established by an Act of Congress and added to the National Park System in 1978. This section of the New River is roughly 50 miles in length and is located between Hinton and Fayetteville, West Virginia (Appendix A). The river flows morthwest through a 1000 foot gorge that has a band of sandstone located near the rim and is densely forested with fifty year old hardwoods throughout. Railroad lines serve as principle access to the area and run the length of the river. Coal, freight, and passenger trains are frequently seen. Several cut-sandstone building foundations can be seen from the river and serve as constant reminders of the numerous abandoned mining towns which once boomed there in the early 1900's.

Commercial whitwater rafting, private boating, and fishing are the major recreational uses of the New River. According to National Park Service estimates, the section of river between Thurmond and Fayette Station (the lower 15 miles of th National River) was boated by more than 83,000 people in 1983. Visitor use of the river varies by season and day of the week. Approximately 72% of all boating use occurs between Memorial Day and Labor Day. In 1983, the weekend day average of commercial river use in the Gorge was 1,026 people/day while the weekday average was 187 people/day.

This study was concerned with commercial raft trips run only on the lower Gorge section of the New River. This section is often described as "some of the best whitewater in the East". Commercially runnable volumes range from about 1000 to 20,000 cubic feet per second and several class IV and V rapids are found in the lower Gorge (based on the rapid classification of I-VI). Over 20 commercial outfitters are licensed by the State to operate whitewater rafting trips on the New River. Several of the smaller companies run but one or two trips on the weekends while several of the larger companies regularly run six or more trips per weekend day.

A typical one day trip in the lower Gorge may range in length from 4 to 8 hours, depending upon the water level and individual company. Likewise, the number of customers on a given trip may range from 20 to 50 people, depending upon the company and day of the week. Much smaller and larger trips are not unusual. On busy Saturdays during the period of peak summer use, it is common to have several trips within sight both upstream and downstream. Trips launch continuously on weekend days at Thurmond from about 8 am until noon. Customers are attracted from all surrounding states.

Study Population and Sampling

The study population included two groups, the comercial river guides and their customers on the New River. More specifically, the population consisted of the guides and guests of Mountain River Tours Inc., a large river outfitter which graciously consented to participate in the study. Only weekend trips in late June to early August were to be included in the

study. Such trips typically had about 40 customers and 5 guides, with one guide per raft. Five guides were selected as test subjects.

Several considerations resulted in restricting study days to the weekends. Weekends receive about five times more use than do weekdays. Restriction to weekends was thus intended to increase the likelihood of full trips (40 customers) from which to sample and to use the same five guides on every day of the study. Weekdays receive far less use, with trips of fewer than 40 customers and five guides common. In addition, knowledge or enjoyment gained from interpretation may interact with the amount of use and crowding on the river. To avoid this potentially confounding variable, only weekend trips were sampled.

The time-frame chosen for the study (late June to early August) was also an attempt to avoid extreme variations in such potentially confounding variables as length of trip due to varying water levels, and weather conditions. Still, both of these factors were measured and recorded on each of the study days.

The same five guides were used throughout the study and were chosen in a stratified random fashion from the entire outfitter's staff of approximately 45 guides. The outfitter's staff of guides was first stratified by experience level. Experienced guides were those with at least one previous season of experience guiding on the New River, and three guides were randomly selected from this stratum[1]. Inexperienced

---

[1]The study investigator was one of the experienced guides included in the study (Guide 1). While the author recognizes that this might bias results, financial constraints and data collection precedures required that he work for an outfitter and be present on study trips. Preliminary analyses suggest that changes in the author's performance following guide training were similar to those of the other study guides.

guides were those with less than one season of guiding experience, and two study guides were randomly selected from this group. Random selection should avoid the threat of selection bias and increase the ability to generalize results to the larger guide population.

On any given Saturday or Sunday during the study, four or five trips were regularly scheduled on each day by the outfitter. Complete random assignment of customers to trips was not possible. Customers were assigned to trips as they made reservations and each trip contained to varying extents related subgroups such as: couples, family units, friend groups and church groups. It was assumed that these subgroups were randomly distributed among all trips. All trips were not filled prior to the start of sampling, and in fact, reservations were taken until the scheduled trip time. This perfunctory procedure for accepting reservations and filling trip space effectively precluded the complete random assignment of individual customers to control or treatment trips. Instead, the five guides participating in the study were randomly assigned to one of the four or five trips scheduled for that day. Occasionally an additional guide was added to a large trip or to replace one of the five study guides if unavailable on that day. Data collected from customers of non-study guides were deleted from the study and not substituted for any of the five study guides.

Research Design

This field experiment utilized a modified Solomon 4 Group Design and a questionnaire for testing study hypotheses. The basic Solomon 4 Group Design can be characterized as follows (Campbell and Stanley, 1976):

```
R   O   X   O       R = Random Assignment of Groups
R   O       O             to Treatment or Control
R       X   O       O = Observation
R           O       X = Treatment
```

The Solomon 4 Group Design increases internal and external validity over other experimental designs since it accounts for both the main effect of testing and the interaction of testing and the treatment (Campbell and Stanley, 1966).The modified Solomon 4 Group Design used in this study can be described as follows:

```
  R   O       O       R = Random Assignment of Groups
  R           O           to Pretest-Posttest or Posttest-only
----- Training ----     O = Observation - Questionnaire
  R   O   X   O       X = Treatment - Guides with Training
  R       X   O
```

Both designs call for customers on a given trip to be either pretested and posttested or posttested only by the questionnaire. On control days, guides provided their customers with a "normal" trip. Following data collections of all control trips, the staff of guides were given a training session by the National Park Service. This training session was the intervention or treatment, for it was a unique educational experience for all the guides with the specific intent and purpose of changing their "normal" trip. Trips that were tested following the training intervention

were considered treatment days, and data were collected in the same manner as on control days. The modified Solomon 4 Group design simply called for treatment trips to chronologically follow the presentation of the guide training program. Customer selection procedures were not altered, and guides were assumed to be the same in control and treatment groups in all respects except the training effect.

Use of the Solomon 4 Group experimental design calls for complete random assignment of subjects to either control or treatment groups. This extent of control was not possible due to the field nature of the study and was one of numerous reasons for the modification of the design. It was deemed necessary to collect all control trip data before the training intervention and all treatment trip data following training. Since the same five guides were used for both control and treatment trips, it was not possible to instruct them to "turn on" the interpretive message one day and "turn it off" the next. Guides are expected to provide their customers with the best trip possible and the training session was intended to improve the trips they offered. Neither was it the intention of the investigator to involve guides in the study to such an extent that it might bias results. Asking guides to discern between a "normal" trip and a treatment trip was considered impractical and subject to bias. The intention in modifying the design was to record what was learned by the customers without purposeful involvement of the guides (and bias) which might have resulted in having them "play dumb" on control days or overreacting on treatment days to "look good".

The same five guides were used throughout the study and on both control and treatment trips in an effort to reduce the difficulty in

interpreting effects due to potentially large individual differences among the guides versus the effects due to treatment. It was assumed that individual differences in each guide's interpretive effectiveness, and in turn the variance in customer scores, would be minimized across control and treatment groups by using the same guides. Randomly assigning different trained and untrained guides to treatment and control days would have required a guide sample size that was beyond the financial, manpower, and logistical constraints of this study. The confusing effect of potentially large individual differences and scores due to different guides was judged to be a greater threat to validity than the potential effect of selection bias and nonequivalence of groups due to the modification of the Solomon 4 Group design. In addition, the selection and assignment of customers to trips and thus to control or treatment groups was made with no control or forethought by the outfitter or investigator. It is acknowledged, however, that early summer visitors may still somehow have been different from late summer visitors. Therefore, several limited ex post facto analyses were conducted to gain evidence on the extent of differences between the two groups. Results of these tests will be discussed under data analysis issues.

Sample Size

Each configuration of observation and/or treatment in the design was initially to be tested over the duration of three one day river trips. There were to be 6 control days and 6 treatment days. These study days were limited to weekend days (Saturday and Sunday) in late June, July and

early August. Only one study trip was to be sampled on any given study day. Control trips were to be randomly selected from all trips on the six days of the first three study weekends. Three of these days were randomly assigned to the posttest-only condition and the other three to the pretest-posttest condition. The guides were then to receive National Park Service training and the next three weekends were to be treatment days. On these treatment days, the study trips were randomly selected from all trips, and half were assigned to the posttest-only condition and the other half to pretest-posttest condition. The same five guides were to run the control trips and the treatment trips. Assuming trips were completely filled, the 12 study days were to provide an expected total sample size of about 480 subjects (Table 1).

Various circumstances and field constraints contributed to the fact that while 6 control days were tested, only 4 days following guide training were likewise tested. Customer cancellations, guide scheduling conflicts, and unexpected rain (which resulted in switching scheduled trips to another nearby river) were the primary causes for the inability to maintain equal sizes of the control and treatment groups. The number of customers tested also varied across guides. These constraints and circumstances resulted in a much different sample size than initially expected (Table 2).

The initial 6 weekend sampling period was lengthened considerably in an attempt to collect data on 6 control days and 6 treatment days. Data collection began the last week of June as planned, and then was extended through the last weekend of August and the end of the New River boating season. Various circumstances still limited data collection to 4

TABLE 1

Planned Sample Size - Distribution of Customers
by Experimental Group and Guide

# Customers Tested	Pre-test		Post-test	# Customers/ Guide	# Trips Tested
120	0		0	24/guide	3
120			0	24/guide	3
240 CONTROL GROUP Customers				48 Customers/Guide	6
----------TRAINING----------					
120	0	X	0	24/guide	3
120		X	0	24/guide	3
240 TREATMENT GROUP Customers				48 Customers/Guide	6

Based on the assumption and testing of twelve fully booked trips of 40 customers/trip and five study guides/trip.

TABLE 2

Actual Sample Size - Final Distribution of Customers
Experimental Group and Guide

# Customers Tested	Pre-test	Post-test	# Customers/guide by Study Guide 1	2	3	4	5	#Trips Tested
85	0	0	22	21	15	22	5	3
86		0	23	21	21	14	7	3
171 CONTROL GROUP Customers			45	42	36	36	12	
			Customers/Guide					
----------Training----------								
55	0	X 0	10	13	16	10	6	2
53		X 0	14	7	17	7	8	2
108 TREATMENT GROUP Customers			24	20	33	17	14	10
			Customers/Guide					Total Trips Tested

Based on the testing of 345 customers of which 66 belonging to nonstudy guides were deleted, resulting in a total sample size of 279 customers.

treatment days. Unusually rainy late summer weather often caused trips to be switched to the Gauley River, as a matter of company policy. Customer cancellations and last minute no-shows sometimes resulted in the cancellation of the entire trip, while other times it was combined with another trip. This often meant study guides were reassigned to other trips or tasks and prevented sampling on that day. Sampling days that were cancelled for the above reasons were rescheduled for the next weekend. Guide training was thereby delayed until August 4, 1984, instead of mid-July.

Cancellations sometimes reduced the trip size to 33 or fewer customers. This situation meant one guide was "bumped" from the trip, and this reduced the number of guides on the trip from 5 to 4. Occasionally a scheduling conflict resulted in the removal of one of the 5 study guides from the trip and replacement by another guide. Since each guide (whether trained or untrained) independently conducted his own raft, no experimental problems were created by not having all five study guides present on each trip sampled. However, all data collected from customers in a non-study guide's raft were deleted from the analyses, resulting in the loss of 66 questionnaires from the total sample size collected. Reduced size trips and the additional scheduling of non-study guides contributed to the variation in sample size across guides and across the posttest-only and pretest-posttest groups.

Guide Training Program - The Treatment

The guide training program consisted of a one-evening information and education seminar conducted by three interpreters of the National Park

Service at Mountain River Tours headquarters. To increase generalizability to realistic staff training conditions, the training was not highly individualized in nature. Rather, the format was one that the National Park Service had used in the past and could use in the future with larger groups in an off-site setting and seminar approach. The entire Mountain River Tours staff was invited to attend and approximately 20 guides (including the study guides) were present. The determination of the training programs content was derived from the New River visitor survey results (Roggenbuck and Bange, 1983) and represents the information and interpretation New River boaters indicated as desirable. General topics included the natural and cultural history in and on the river, the national significance of the New River Gorge National River, the history and purpose of the National Park Service, and information available at the visitors centers. The emphasis of this training was to present information to the guides to increase their knowledge base and to encourage their interpretive presentation of this information to their customers.

The following guide training content outline was developed in cooperation with the interpretive staff of the Park Service:

- Goals and Objectives of the Training Session
- National Park Service historical background
- New River Gorge National River historical background, present state, future goals and projects (oral presentation with maps).
- Cultural History of railroad and mining towns in the New River Gorge (Historical slide and photo presentation).

- Coke Production in New River Gorge (movie).
- Oral History Program, a presentation by local citizen and former inhabitant of a New River Gorge mining town (stories and about life and the way it was living and working in the Gorge in early 1900's).
- Natural History of the New River (scripted slide show from headwaters of the New River to its confluence with the Gauley River).
- Birds commonly seen in the Gorge (slide presentation and bird calls).
- Conclusions - review of goals & objectives met durint training session.

The guide training session was about four hours in length with several breaks incorporated into that time frame. Information was presented in a multimedia format utilizing oral, slide, movie, diagrammatic and show-and-tell presentations. Questions were encouraged and further sources of information suggested to augment answers to those questions.

Questionnaire

Pretest and Posttest Forms

The pretrip questionnaire consisted solely of 13 multiple choice questions developed to assess the customer's knowledge of the National Park Service, the New River Gorge National River, the significance of the New River Gorge, and the natural and cultural history of the river

(Appendix B). The identical 13 questions comprised the first page of the posttrip questionnaire. The second page of the posttrip questionnaire asked 10 additional questions. Questions one through seven on the second page used a seven-point Likert format to determine such things as how much their guide talked about the topics of the previous 13 questions, whether they enjoyed listening to those sorts of things, how they rated the trip overall, and whether they intended to visit a New River Gorge visitor center (Appendix C). These single item measures were developed and pilot tested under consultation and review of river recreation researchers and National Park Service staff. Question 8 asked customers whether it was their first, second, third, or fourth or more time rafting the New River. Questions 9 and 10 were open ended questions asking what the customers liked or disliked about their trip.

Administrative Procedures

The pretest questionnaire was administered to all customers on the 45-55 minute bus ride to the put-in on pretest survey days. Customers were asked to initial the form as a means of identification. Upon arrival at the put-in, the customers separated into smaller groups of 7-9 and chose a raft and guide. The random assignment of customers to guides was not realistically possible nor attempted.

All study trips were administered the posttest questionnaire immediately following the trip on the 20-25 minute bus ride back to the outfitter's base camp. This posttest questionnaire included the same 13 multiple choice questions found on the pretest, as well as a second page

of questions concerning their trip enjoyment, amount of interpretation presented on the trip, customer's river running experience, their guide's name, and things they liked or disliked about their trip. The posttest questionnaire was also initialed by the customers and collected as they exited the bus.

Identification of the questionnaire by initials or name was generally a successful means of matching pretests and posttests. Eight questionnaires were not identified and could not be matched, and five did not identify their guide. These questionnaires were deleted from the data set.

The protocol for administration of the questionnaire was standardized through use of a script (see Appendices D and E). Customers were asked by the study investigator to answer all the questions and to choose the best answer if they were unsure of the correct response. The script tried to emphasize the idea that the surveys were not being "graded", that their responses were confidential, and that "cheating" was not necessary. Still, some couples answered the questionnaire as a single unit, and some groups actively traded answers. Thus, some degree of cooperative testing was noted.

Administration of the posttest questionnaire to control groups was particularly difficult and frustrating for the customers since untrained guides had apparently seldom covered areas of the questionnaire's subject matter during the trip. In addition, half of these groups had to take the same test twice. Answers to all the questions were discussed with the customers following collection of all posttests to help overcome this frustration. However, questionnaires of five individuals, who had

completed all thirteen questions on the pretest but failed to answer five or more knowledge questions on the posttest, were deleted from the study since their potential score did not include the benefit of guessing on all thirteen questions.

A wide variation in the ability of customers to read and answer the questions was noted. Some took much longer than others, some were much more cooperative than others. These effects, however, seemed well distributed throughout all test days.

Instrument Development - Validity Issues

Content validity refers to the degree to which the test items represent the content domain to be measured (Anastasi, 1968). Knowledge questions 1-13 on the pretest and posttest instruments were generated in a manner so as to help assure content validity. First an attempt was made at exhaustively defining the domain of knowledge concerning the selected topics of the guide training program. A list of potential items was then generated and submitted for scrutiny by Virginia Tech recreation researchers and committee members, Mountain River Tours staff, and interpreters for the National Park Service at the New River Gorge. Factual accuracy, grammatical correctness, appropriateness of level of difficulty, readability, and most importantly, the representativeness of the content, were criteria for the reduction and elimination of items from the pool.

Asking too many questions and overburdening the customer was a major concern that Wagar (1976) warned against. This was one reason for

limiting the length of the final pretest questionnaire to one page and the posttest questionnaire to two pages. The need for a noncumbersome and easily administered instrument was also necessitated by constraints inherent to the field setting in which data were collected. The 20-25 minute bus ride from the river to the outfitters headquarters created a time frame restriction for instrument administration. It was also necessary for customers to be able to read and write on a moving bus on bumpy secondary roads while frequently still dripping wet.

A list of 17 knowledge questions was pilot tested with 48 customers on a weekend trip in early June, 1984. An item analysis of responses to the 17 questions helped identify poor questions and further reduce the pool of items. The final 13 items on the questionnaire were then reviewed by the "experts" most knowledgeable of the content domain, namely the chief interpreter and his assistants at the Park Service, and judged to be highly representative of the domain of knowledge being considered, and therefore valid.

Instrument Development - Reliability Issues

Reliability refers to the consistency with which an instrument measures that which it is intended to measure. Cronbach's Alpha, which yields the arithmetic mean of all possible split-half reliability coefficients, was used to evaluate the reliability of the 13 knowledge items. All data were coded, and then responses were scored by computer as either correct or incorrect responses using a standardized key to remove the threat to reliability due to inconsistencies in scoring. The

Kuder-Richardson Formula (KR-20) was used since responses were coded dichotomously.

Cronbach's Alpha coefficients equalled .608 for the knowledge posttest scores and .439 for the pretest scores. Alpha coefficients for pretest scores were expected to be lower since customers were by and large guessing at most answers. Considerably less guessing occurred with posttest scores following training and thereby increased correlations among items. The postest reliability of .608 is considered adequate for research concerning assessment of a group characteristic (Nunnally, 1967).

DATA ANALYSIS

Scoring the Questionnaire

Knowledge questions (1-13) on the pretest and posttest questionnaires were scored as correct or incorrect by computer. Correct respones on the items of the pretest were then summed to create a knowledge test score called the prescore. Posttest item responses were scored in the same way to give an overall postscore. Missing data was interpreted as an incorrect response. Simple gain scores were calculated as the difference between an individual's postscore and prescore.

Group Equivalence

Since the procedure used in this study to select and assign customers to experimental conditions did not guarantee equivalence of control and treatment groups, it was important to investigate this issue more fully. With random assignment, treatment groups are considered to be initially equivalent (Cook and Campbell, 1979). However, without complete randomization, estimates of the effects of treatment may be biased. A selection procedure which is not formally random, however, may still be sufficiently haphazard so that it is "random in effect" (Lord, 1963). Cook and Campbell (1979) describe a two-step procedure for the use of ANOVA employing both pretest and posttest scores which addresses this issue. The pretest scores are first analyzed for statistical significance of differences. Mean pretest scores between control and treatment groups

which do not differ significantly suggest equivalence of the groups. Though such lack of differences is not sufficient to guarantee equivalence, it can be considered additional evidence of equivalence of the two groups. This is in turn justifies the second step of data analysis where the ANOVA model is applied to the posttest scores to determine treatment effects.

Following this procedure, the mean prescores on the knowledge questions were examined for differences between control and treatment groups. Due to the unequal sizes of the comparative groups, the SAS General Linear Model (GLM) was utilized to identify significant differences between the least square means of the prescores. Results indicate that differences in the least square means were nonsignificant (Table 3). Frequencies of correct and incorrect customer responses to each pretest question (1-13) are summarized in Table 4.

Differences in prescore variance were also tested and found to be nonsignificant (Table 3). Procedures followed for variance testing are described by Snedecor and Cochran (1980). The noted similarity of means and variances of prescores suggests that control and treatment groups had similar distributions of prescores.

Cook and Campbell (1979) argue that equivalence of the groups cannot be adequately determined on the basis of pretest scores alone. For this reason, the experience level of the customers was also examined for differences between control and treatment groups. Warden and Raile (1984) found that visitors to a historic site who had made previous visits exhibited significantly higher knowledge scores than first time visitors. Experience level of customers on the New River could thereby

TABLE 3

Comparison of the Prescore LS Means and Variance of Control and Treatment Groups

Dependent Variable: Prescore Means

Source	df	S.S.	F	PR > F
Model	1	.06229	.01	.9037
Error	137	580.35499		

Training	Prescore LS Means	Standard Error LS Means	N	Variance
Before-Control	4.048	.225	84	4.4796
After-Treatment	4.091	.278	55	3.8619

Dependent Variable: Prescore Variances

H_o: variances equal F = 4.4796/3.8619 = 1.159 df 83,54

The null hypothesis is rejected at the .05 level of significance if F > 1.69 (5% two-tailed value). Since F = 1.159 which is less than the Table value of 1.69, the null hypothesis is rejected.

TABLE 4

Frequencies (%) of Pretest Responses to Knowledge Questions (1-13) Before and After Guide Training

Question[1]	Before Guide Training				After Guide Training			
	Correct		Incorrect		Correct		Incorrect	
	N	%	N	%	N	%	N	%
1	19	22.3	66	77.7	16	29.1	39	70.9
2	19	22.3	66	77.7	9	16.4	46	83.6
3	23	27.1	62	72.9	19	34.6	36	65.4
4	42	49.4	43	50.6	31	56.4	24	43.6
5	17	20	68	80	8	14.6	47	85.4
6	20	23.5	65	76.5	15	27.3	40	72.7
7	31	36.5	54	63.5	28	59.9	27	49.1
8	25	29.4	60	70.6	15	27.3	40	72.7
9	17	20	68	80	10	18.2	45	81.8
10	55	64.7	30	35.3	36	65.4	19	34.6
11	31	36.5	54	63.5	21	38.2	34	61.8
12	20	23.5	65	76.5	11	20	44	80
13	21	24.7	64	75.3	6	10.9	49	89.1

[1]Questions 1-13 are listed in Appendix B.

have a potentially serious effect on their knowledge prescores and postscores. Frequencies of customer experience levels both before and after training do not indicate any obvious differences among the groups (Table 5). Results of a statistical comparison of the least square means of the control and treatment groups indicate nonsignificant differences at the .05 level. Differences in the variance of customer response to experience level were also nonsignificant (Table 6).

The fact that differences in the least square mean and variance of both prescores and customer experience levels were nonsignificant strongly suggests that the control and treatment groups have similar prescore and experience level distributions. The strong similarity of these distributions further suggests that the control and treatment groups were initially equivalent and therefore were assumed to be so. The potential bias of selection effects on treatment outcomes was also assumed to be minimized.

ANOVA of Postscores and Trip Variables

The statistical technique of ANOVA was used in hypothesis testing to determine the presence of treatment effects (Cook and Campbell, 1979). In particular, the SAS General Linear Model procedure (GLM) was utilized due to the unequal sizes of the control and treatment groups. It is acknowledged that Likert-scaled questions are generally ordinal in nature and that the use of means as a description of central tendency generally requires interval data, such as the knowledge postscores (Mason and Bramble, 1978). Research, however, has shown that the ANOVA is robust with respect to violations of the normality assumptions and use of ordinal

TABLE 5

Frequencies of Visitors' Experience Level Rafting the New River Before and After Guide Training

Customer Experience Level		Before Training (N)	Before Training (%)	After Training (N)	After Training (%)
1	first trip	101	61.2	62	59.1
2	second trip	36	21.8	22	20.9
3	third trip	14	8.5	11	10.5
4	fourth or more trips	14	8.5	10	9.5
	Total N	165	100%	105	100%

TABLE 6

Comparison of Customer Experience Level
of Control and Treatment Groups

Dependent Variable: Mean Experience Level

Source	df	S.S.	F	PR > F
Model	1	.24935	.26	.608
Error	268	253.75065		

Training	Experience LS Means	Standard Error LS Means	N*	Variance*
Before (Control)	1.6424	.0757	82	.80593
After (Treatment)	1.7048	.0950	54	.89902

* Using data from Pre- and Posttested Groups only.

Dependent Variable: Variance in Visitors' Experience Level

H_o: variances equal F = .89902/.80593 = 1.1155 df 53,81

The null hypothesis is rejected at the .05 level of significance if F> 1.68 (5% two-tailed test). Since F = 1.115 which is less than the Table value of 1.68, the null hypothesis is not rejected.

gain score analysis was another statistical method used to determine if knowledge scores had increased following training.

Tests of the Effect of Weather and Water level

Weather and water levels of the river were two potentially confounding variables which might have had significant effects on the study's dependent variable means. The range of water levels which are commercially runnable can result in trips varying in length from four to eight or more hours. High water levels result in much shorter trips which permit the guide less time to talk to the visitors. The effect of variable weather conditions on trip enjoyment or visitors' interest in interpretation was unknown but considered potentially significant. These two variables were therefore included in the ANOVA tests of study hypotheses. Weather was operationalized as a five-category independent variable (1 = cold and rainy to 5 = warm and sunny), while water levels were recorded from river gauge readings in cubic feet per second volumes and included in the model as a covariate. Water level was never found to be important, and weather was significant ($p<.0455$) in explaining differences in least square means of only the customer intentions to visit a New River Gorge visitor center.

Spearman Correlation Coefficients

Spearman correlation coefficients were employed in examining the relationships among three posttest dependent variable responses. The

amount of interpretation presented on the river was correlated with visitors' overall trip ratings. The overall trip rating was also correlated with the extent to which guides discussed natural and cultural history.

RESULTS

Data Overview

Frequencies of posttest customer responses to the thirteen knowledge questions and the Likert-scaled questions (Appendix C) suggest many changes from before training to after training (Table 7 and 8). The number and percentage of correct responses to each knowledge question increased following training. When customers were asked how much their guide talked about the topics mentioned in the previous knowledge questions, a greater percentage of customers answered "alot, all the time" following training than before training, suggesting an increase in the amount of interpretation presented. Also, following training, a larger percentage of customers indicated that listening to and learning about the sorts of things mentioned in the knowledge questions added to the enjoyment of their trip. When asked how they rated their trip overall, an apparently greater percentage of visitors indicated their trip was "excellent" or "perfect" following training than before training. However, frequencies of response regarding intentions to visit a National Park Service visitor center did not seem to change following training.

Table 5 reports visitors' rafting experience level by the number of trips they had taken on the New River. There was very little difference in the frequency of customers' experience levels before and after training.

TABLE 7

Frequencies (%) of Posttest Responses to Knowledge Questions (1-13) Before and After Guide Training

Question[1]	Before Guide Training				After Guide Training			
	Correct		Incorrect		Correct		Incorrect	
	N	%	N	%	N	%	N	%
1	57	33.3	114	66.7	81	75	27	25
2	53	31	118	69	60	55.6	48	44.4
3	67	39.2	104	60.8	94	87	14	13
4	107	62.6	64	37.4	82	75.9	26	24.1
5	33	19.3	138	80.7	30	27.8	78	72.2
6	69	40.3	102	59.7	57	52.8	51	47.2
7	101	59.1	70	40.9	83	76.8	25	23.2
8	49	28.6	122	71.4	78	72.2	30	27.8
9	76	44.4	95	55.6	54	50	54	50
10	111	64.9	60	35.1	89	82.4	19	17.6
11	79	46.2	92	53.8	88	81.5	20	18.5
12	79	46.2	92	53.8	73	67.6	35	32.4
13	57	33.3	114	66.7	55	50.9	53	49.1

[1]Questions 1-13 are listed in Appendix B.

TABLE 8

Frequences (%) of Visitors' Posttest Responses on Trip Variables Before and After Guide Training

Question	Response Scale 1	2	3	4	5	6	7
1. How much would you say your guide talked about the topics mentioned in the previous questions?							
Before	.6	.6	1.2	19.4	27.3	38.8	12.1
After	0	0	0	5.7	15.2	44.8	34.3
	Never, not a word			Sometimes		Alot, all the time	
2. Does listening to and learning about these sorts of things add to or detract from the enjoyment of your trip?							
Before	0	0	.6	8.5	7.9	24.8	58.2
After	0	0	0	4.8	4.8	20.0	70.5
	Detract from			Neutral			Add to
3. Overall, how would you rate your trip?							
Before	0	0	0	5.5	24.8	44.2	25.4
After	0	0	.9	1.0	13.3	58.1	26.7
	Unacceptable	Poor	Fair	Good	Very Good	Excellent	Perfect
4. To what extent did your guides' discussion of natural and cultural history contribute to your overall rating of the trip?							
Before	2.4	1.8	5.4	16.9	28.9	27.7	16.9
After	0	3.8	1.9	6.7	22.9	37.1	27.6
	Not at all						Very Much

TABLE 8 (cont.)

Question	Response Scale 1	2	3	4	5	6	7
5. Do you wish he had talked about New River Gorge natural and cultural history more or less?							
Before	.6	0	2.4	51.2	23.8	11.6	10.4
After	.9	1.0	1.9	63.8	14.3	10.5	7.6
	Less			About Right			More
7. Do you intend to visit any of the New River Gorge Visitor Centers?							
Before	7.9	3.6	2.4	39.4	17.0	9.1	20.6
After	13.3	.9	3.8	46.7	9.6	16.7	19.0
	NO!			Maybe			YES!

Sample Size Before 164-166 - Control
After 104-105 - Treatment

Hypothesis Tests[2]

This section presents the results of the testing of the study hypotheses.

Hypothesis One: Following guide training, the amount of interpretation presented on the river will increase.

The operational measure of this hypothesis was the following Likert-scaled question: How much would you say your guide talked about the topics mentioned in the previous questions? The amount of interpretation presented on the river was therefore operationalized as the relative amount of time customers reported their guide talked about the topics mentioned in the thirteen multiple choice knowledge questions. For statistical purposes, the null hypothesis tested was: there will be no significant difference in the amount of interpretation presented on the river before and after training.

Significant differences in the mean visitor ratings of the amount of interpretation presented on the trip before and after training were due to the independent variables of guide and group (Table 9).

---

[2]The author realizes that from a theoretical standpoint, a statistical test can only be conducted on a null hypothesis. Hypotheses are first stated in non-null directional form based on past empirical and theoretical discussion found in the literature. Statistical tests were, however, performed on the null non- directional hypotheses. If the null hypothesis was rejected, the direction of differences was determined by examination of the least square means.

TABLE 9

Test for Differences in Visitors' Ratings of the Amount of Interpretation Presented on River by Guide and Group (Experimental Condition)

Source	df	F value	PR > F
GUIDE	4	8.02	.0001
- Exp. vs. Inexp.	1	7.40	.0070
GROUP	3	15.71	.0001
- Pretest	1	3.79	.0527
- Treatment	1	40.08	.0001
- Interaction	1	0.65	.4210
GUIDE * GROUP	12	0.97	.4775
Error	250		
Total	269		

Group	LS Mean[1]	Std. Err. LS Mean
Pretest-Posttest Treatment	6.128	0.132
Pretest-Posttest Control	5.429	0.118
Posttest-Only Treatment	5.984	0.143
Posttest-Only Control	5.081	0.111

Guide	LS Mean	Std. Err. LS Mean
1 - Experienced	6.120	0.119
2 - Experienced	5.979	0.142
3 - Inexperienced	5.644	0.113
4 - Inexperienced	5.260	0.138
5 - Experienced	5.275	0.184

[1]Amount of interpretation presented is based on seven-point Likert scale.

Results indicated significant differences ($p<.0001$, $F=15.71$) among the least square means of the four possible combinations of pretest-posttest, posttest-only, control and treatment experimental conditions which are referred to as the group variables. The variance due to the group effect was partitioned into components due to pretesting, treatment, and the interaction effect of pretesting and the treatment. Of these three, only the effect due to treatment was significant ($p<.0001$, $F=40.08$). This indicated significant differences between the treatment and control groups. Examinations of group least square means indicated an increase in mean response following training (Table 9). The differences in mean response due to the pretesting effect approached significance at the .05 level ($p<.0527$, $F=3.79$) and an increase in the pretested group means over the groups not pretested was noted (Table 9).

The interaction effect of the pretest and treatment, which potentially could have increased those posttest scores beyond the effect due to treatment, proved nonsignificant ($p<.4210$, $F=0.65$). This served as justification for combining data of groups for a comparison of untrained control group versus a trained treatment group. Results from this analysis indicated a significant difference ($p<.0001$, $F=38.38$) in the time guides spent talking about the topics of interest before and after training (Table 10). Examination of the least square means indicated an increase following training (Table 10).

The significant difference ($p<.0001$, $F=8.02$, Table 9) in visitors' rating of the amount of interpretation presented due to the guides was partially explained by the experience level of the guides ($p<.007$, $F=7.40$, Table 9). Two of the three experienced guides were rated as presenting

TABLE 10

Test for Differences in Visitors' Ratings of the Amount of Interpretation Presented on River by Guide and Training

Source	df	F value	PR > F
GUIDE	4	7.42	.0001
PRETEST	1	5.34	.0217
TRAINING	1	38.38	.0001
PRE*TRAINING	1	2.48	.1167
GUIDE*TRAINING	4	0.86	.4888
Error	258		
Total	269		

Training	LS Mean	Std. Err. LS Mean
Before	5.277	0.080
After	6.042	0.094

Pretest	LS Mean	Std. Err. LS Mean
Yes	5.795	0.080
No	5.524	0.084

	Guide	Training	LS Mean	Std. Err. LS Mean
1	Experienced	Before	5.94	.141
		After	6.30	.189
2	Experienced	Before	5.51	.148
		After	6.37	.221
3	Inexperienced	Before	5.23	.156
		After	6.07	.163
4	Inexperienced	Before	4.84	.154
		After	5.76	.224
5	Experienced	Before	4.87	.267
		After	5.72	.247

more interpretation than the two inexperienced guides (Table 9). It should be noted, however, that the amount of information presented on the river increased following training for all guides (Table 10).

The model using trained versus untrained guide effects (Table 10) further indicated significant pretest effects ($p<.0217$, $F=5.34$). Pretested customers reported greater least square means than customers that were posttested-only (Table 10). Perhaps the pretest was viewed by customers as an additional amount of interpretation and was attributed to the guide following the trip. Another possibility is that the pretest fostered interest in learning more about the river and prompted customers to ask their guide for more information. The significant pretest effect suggests examination of the least square means of the posttest-only groups as a more accurate and realistic estimate of the treatment effects.

Hypothesis Two: Following guide training, visitors' knowledge of the National Park Service, the national significance of the New River Gorge National River, and the natural and cultural history of the river will increase.

This hypothesis was tested by use of two separate analyses. The first examined differences before and after training in the visitor's mean postscores on the thirteen multiple choice knowledge questions. The second analysis examined differences in the visitors' mean gain scores. If the replication of results from two analyses which revealed significant treatment effects this would increase confidence in the tested hypothesis. Therefore, the first operational measure of this hypothesis was the mean visitor postscores on the thirteen item knowledge index. The

null hypothesis first tested was: there will be no significant difference in visitors' knowledge (postscores) following training.

Results indicate significant differences in the means of posttest knowledge scores due to the group and the guide by group interaction term (Table 11). The group was highly significant ($p<.0001$, $F=46.89$) in explaining variation in the postscore least square means. The group effect was again partitioned into components due to pretesting, treatment, and interaction effects. Treatment effects were highly significant ($p<.0001$, $F=124.48$) and examination of postscore means indicated increases in the treatment groups (Table 11). The pretest also accounted for significant ($p<.007$, $F=11.80$) although much smaller, variation in the postscore means (Table 11). Pretested group postscores were greater than posttest-only grop postscores by approximately one correct response on the thirteen item knowledge index (Table 11).

The interaction effect of the pretest and treatment was nonsignificant ($p<.8821$, $F=0.02$). This indicated that the pretest did not uniquely interact with the treatment to increase postscores (Table 11). Instead, the pretest effect was consistent across both control and treatment groups. A comparison of control group versus treatment group postscore means indicated significant differences ($p<.0001$, $F=126.53$) in scores following training (Table 12). Examination of the least square means of postscores indicated an increase in scores following guide training of approximately three correct responses on the thirteen item index (Table 12).

The significant interaction effect of guide by group indicated that guides were not consistent across groups. This is further discussed under hypothesis five.

TABLE 11

Test for Differences in Visitors' Posttest Knowledge Scores[1] by Guide and Group (Experimental Condition)

Source	df	F value	PR > F
GUIDE	4	1.80	.1298
- Exp. vs. Inexp.	1	6.26	.0130
GROUP	3	46.80	.0001
- Pretest	1	11.80	.0007
- Treatment	1	124.48	.0001
- Interaction	1	0.02	.8821
GUIDE * GROUP	12	2.60	.0028
Error	259		
Total	278		

Group	LS Mean	Std. Err. LS Mean
Pretest-Posttest Treament	8.965	0.294
Pretest-Posttest Control	5.831	0.263
Posttest-Only Treatment	7.972	0.304
Posttest-Only Control	4.919	0.245

Guide	LS Mean	Std. Err. LS Mean
1 - Experienced	6.759	0.263
2 - Experienced	6.914	0.289
3 - Inexperienced	7.160	0.250
4 - Inexperienced	7.508	0.309
5 - Experienced	6.268	0.411

[1]Based on number of correct responses to 13 multiple choice questions.

TABLE 12

Test for Differences in Visitors' Posttest Knowledge Scores by Guide and Training

Source	df	F value	PR > F
GUIDE	4	2.27	.0622
PRETEST	1	7.62	.0062
TRAINING	1	126.53	.0001
PRETEST*TRAINING	1	0.24	.6232
GUIDE*TRAINING	4	4.03	.0035
Error	267		
Total	278		

Training	LS Mean	Std. Err. LS Mean
Before	5.402	0.180
After	8.511	0.210

Pretest	LS Mean	Std. Err. LS Mean
No	6.598	0.190
Yes	7.316	0.189

Guide	Training	LS Mean	Std. Err. LS Mean
1 - Experienced	Before	5.632	0.311
	After	7.924	0.428
2 - Experienced	Before	5.833	0.322
	After	8.161	0.471
3 - Inexperienced	Before	5.515	0.349
	After	8.888	0.364
4 - Inexperienced	Before	5.045	0.350
	After	10.183	0.508
5 - Experienced	Before	4.987	0.604
	After	7.399	0.599

Cook and Campbell (1979) have also suggested an analysis of variance using gain scores to examine the difference, or change, in performance from the pretest to the posttest. The assumption is that the treatment effect would lead to a greater change in the treatment group than in the control group. The null hypothesis tested was: there will be no significant difference in mean change of visitors' knowledge (gain scores) before and after training.

Results indicated significant differences ($p<.0001$, $F=46.95$) in gain scores due to the independent variable of group (Table 13). Examination of the least square means indicated that treatment group means were greater than control group means. Visitors tested following guide training showed changes in knowledge scores mean of approximately three more correct responses than did visitors tested before guide training.

Examination of the frequency distribution of postscores, before and after training, suggested an increase in scores across a wide range of visitor knowledge levels (Table 14). This further indicated that increases in knowledge postscores following training were not due to a small group of history buffs with extreme outlier scores.

Hypothesis Three: Following guide training, visitor ratings of the quality of the trip will increase.

The operational measure of this hypothesis was the response to the following seven point Likert-scale question: Overall, how would you rate your trip? The underlying premise of the hypothesis is that a more enjoyable trip will be more highly rated. Part of the larger objective was to see if, following training, increases in the amount of

TABLE 13

Test for Differences in Visitors' Knowledge Gain Scores by Guide and Group

Source	df	F value	PR > F
GUIDE	4	1.24	.2984
- Exp. vs. Inexp.	1	0.51	.4756
GROUP	1	46.95	.0001
- Treatment	1	46.95	.0001
GUIDE*GROUP	4	1.99	.0994
Error	129		
Total	138		

Group	LS Mean	Std. Err. LS Mean
Pretest-Posttest-Treatment	4.902	0.330
Pretest-Posttest-Control	1.863	0.297

Guide	Group	Gain Score LS Mean
1 - Experienced	Treatment	3.200
	Control	2.045
2 - Experienced	Treatment	4.692
	Control	2.143
3 - Inexperienced	Treatment	4.750
	Control	1.643
4 - Inexperienced	Treatment	6.200
	Control	1.682
5 - Experienced	Treatment	5.667
	Control	1.800

TABLE 14

Frequencies (%) of Postscore Values
Before and After Guide Training

Postscores (# of correct responses)	Before Training N	Before Training %	After Training N	After Training %
1	1	1.2	0	0
2	8	9.4	0	0
3	9	10.6	0	0
4	11	12.9	2	3.6
5	9	10.6	3	5.4
6	7	8.2	3	5.5
7	14	16.5	6	10.9
8	12	14.1	6	10.9
9	10	11.8	14	25.5
10	2	2.4	9	16.4
11	2	2.3	4	7.3
12	0	0	5	9.1
13	0	0	3	5.4
	N = 85		N = 55	

interpretation presented on the river were correlated with any potential increases in visitors' ratings of the trip. The null hypothesis tested was: there will be no significant difference in the mean trip rating before versus after guide training.

Significant differences in the mean rating of the trip were due to the independent variables of guide and group, as well as the guide by group interaction term (Table 15). Results indicated that significant differences ($p<.0101$, $F=3.86$) in the mean trip rating were due to the group effect. Of the three partitioned components of the group effect, only the treatment was significant ($p<.0026$, $F=9.28$, Table 15). Examination of group least square means indicated that treatment trips received higher ratings than control trips (Table 15).

Significant differences in the mean rating of the trip were also due to the guides ($p<.0001$, $F=7.32$, Table 15), but this was not explained by their experience level ($p<.5560$, $F=0.35$). The least square means of trip rating were higher for some guides than others (Table 15). The significant ($p<.0078$, $F=2.32$) guide by group interaction effect indicated that differences between groups were not consistent across guides (Table 16). Some guides received consistently high trip ratings while others showed considerable variation in trip ratings across groups (Table 16, Figure 1).

Results of the untrained control group and trained treatment group analyses indicated that differences in the trip rating means were due to the guide, training, and guide by training interaction effects (Table 17). Least square means of the trip rating were greater following guide training than before. The significant guide by training interaction

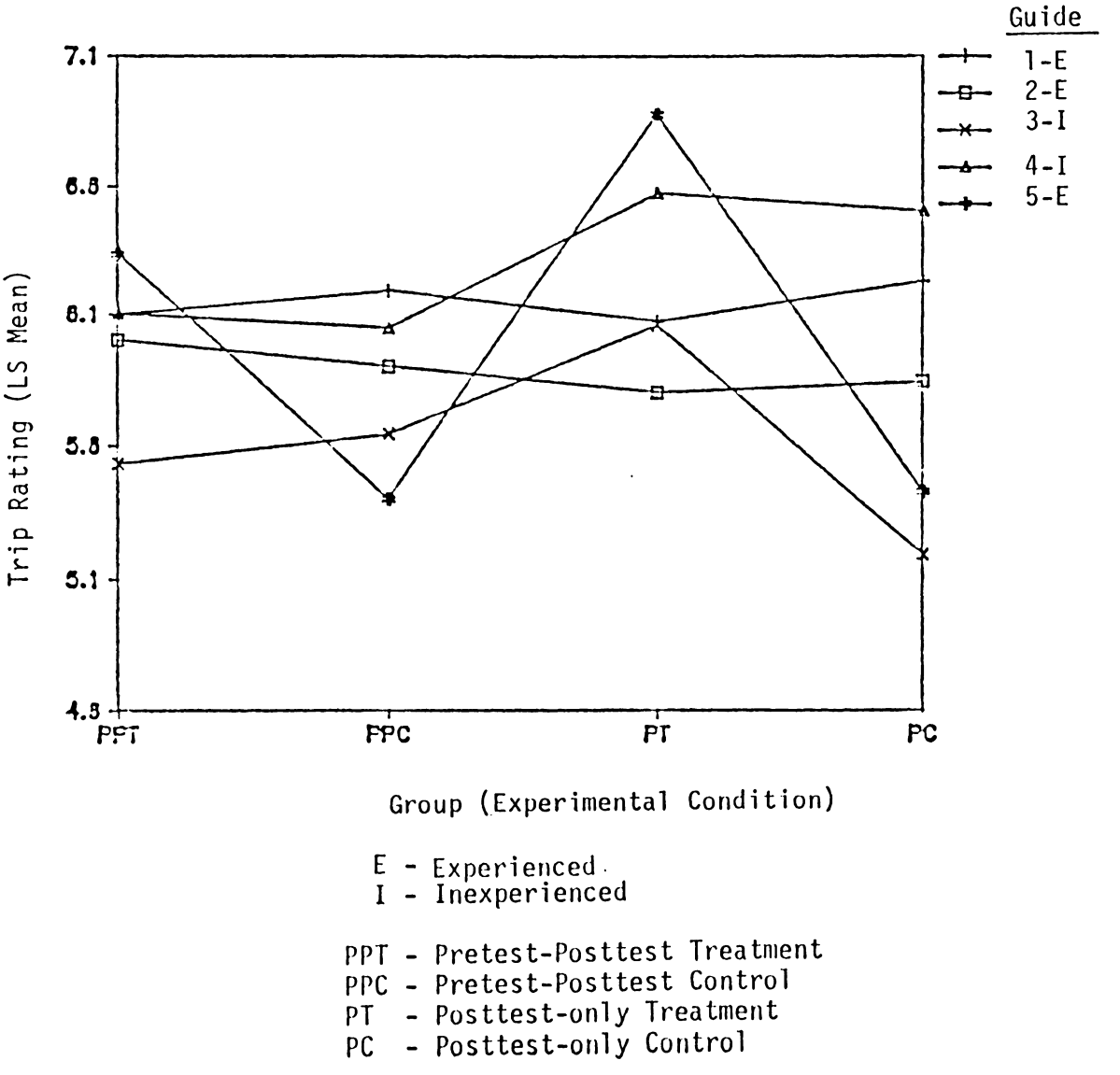


Figure 1. Guide by Group Interaction: Mean trip ratings associated with 5 study guides across 4 experimental conditions.

TABLE 15

Test for Differences in Visitors' Rating of the Trip by Guide and Group

Source	df	F value	PR > F
GUIDE	4	7.32	.0001
- Exp. vs. Inexp.	1	0.35	.5560
GROUP	3	3.86	.0101
- Pretest	1	1.71	.1924
- Treatment	1	9.28	.0026
- Interaction	1	1.66	.1990
GUIDE*GROUP	12	2.32	.0078
Error	250		
Total	269		

Group	LS Mean	Std. Err. LS Mean
Pretest-Posttest Treatment	6.013	0.105
Pretest-Posttest Control	5.836	0.095
Posttest-Only Treatment	6.275	0.114
Posttest-Only Control	5.838	0.088

Guide	LS Mean	Std. Err. LS Mean
1 - Experienced	6.147	0.095
2 - Experienced	5.886	0.113
3 - Inexperienced	5.606	0.091
4 - Inexperienced	6.304	0.110
5 - Experienced	6.009	0.146

TABLE 16

Guide by Group Comparison of LS Means
of Visitors' Trip Rating

Guide	Group	LS Mean
1 - Experienced	Pretest-Posttest Treatment	6.100
	Pretest-Posttest Control	6.190
	Posttest-Only Treatment	6.071
	Posttest-Only Control	6.227
2 - Experienced	Pretest-Posttest Treatment	6.000
	Pretest-Posttest Control	5.900
	Posttest-Only Treatment	5.800
	Posttest-Only Control	5.842
3 - Inexperienced	Pretest-Posttest Treatment	5.533
	Pretest-Posttest Control	5.643
	Posttest-Only Treatment	6.059
	Posttest-Only Control	5.190
4 - Inexperienced	Pretest-Posttest Treatment	6.100
	Pretest-Posttest Control	6.045
	Posttest-Only Treatment	6.571
	Posttest-Only Control	6.500
5 - Experienced	Pretest-Posttest Treatment	6.333
	Pretest-Posttest Control	5.400
	Posttest-Only Treatment	6.875
	Posttest-Only Control	5.426

TABLE 17

Test for Differences in Visitors' Rating of the Trip by Guide and Training

Source	df	F value	PR > F
GUIDE	4	7.14	.0001
PRETEST	1	2.28	.1323
TRAINING	1	12.09	.0006
PRETEST*TRAINING	1	2.28	.1325
GUIDE*TRAINING	4	4.26	.0024
Error	258		
Total	269		

Training	LS Mean	Std. Err. LS Mean
Before	5.818	0.064
After	6.162	0.075

Guide	Training	LS Mean	Std. Err. LS Mean
1 - Experienced	Before	6.209	0.113
	After	6.060	0.151
2 - Experienced	Before	5.872	0.118
	After	6.008	0.177
3 - Inexperienced	Before	5.371	0.125
	After	5.804	0.131
4 - Inexperienced	Before	6.222	0.124
	After	6.319	0.180
5 - Experienced	Before	5.417	0.214
	After	6.622	0.198

effect ($p<.0024$, $F=4.26$) indicated that the increases in means following training were not consistent across guides (Table 17). In fact, four of the five guides showed variable increases in least square means following training while one guide showed a decrease. This suggests that the guide training program helped improve trip ratings for some guides and not for others.

Hypothesis Four: Following guide training, visitors will show increased intentions to visit a New River Gorge visitor center.

The operational measure of this hypothesis was the following Likert-scaled question: Do you intend to visit any of the New River Gorge visitor centers? For statistical purposes, the null hypothesis tested was: there will be no significant difference in visitors' intentions to visit a visitor center before versus after guide training.

Significant differences in the visitors' mean response concerning their intentions to visit a New River Gorge visitor center were due to the guide and weather effects. The group (experimental condition) did not explain any significant variation in the means (Table 18). The treatment did not account for any significant differences in the means ($p<.0787$, $F=3.12$, Table 18).

Significant differences in the means were due to the guide ($p<.0007$, $F=5.02$), but this was not explained by experience level. Examination of the least square means indicated that the customers of some guides stated higher intentions of visiting a visitor center than the customers of other guides (Table 18). The guide by group interaction effect was nonsignificant, indicating differences in the guides were consistent

TABLE 18

Test for Differences in Visitors' Intentions to Visit[1] a New River Gorge Visitor Center by Guide and Group

Source	df	F value	PR > F
GUIDE	4	5.02	.0007
- Exp. vs. Inexp.	1	0.50	.4823
GROUP	3	1.33	.2631
- Pretest	1	0.13	.7209
- Treatment	1	3.12	.0787
- Interaction	1	0.53	.4688
GUIDE*GROUP	12	1.08	.3759
WEATHER	4	2.47	.0455
WATER LEVEL	1	0.93	.3369
Error	245		
Total	269		

Group	LS Mean	Std. Err. LS Mean
Pretest-Posttest Treatment	4.556	.361
Pretest-Posttest Control	4.911	.333
Posttest-Only Treatment	4.151	.424
Posttest-Only Control	5.027	.311

TABLE 18 (Cont.)

Guide	LS Mean	Std. Err. LS Mean
1 - Experienced	5.331	0.238
2 - Experienced	4.424	0.271
3 - Inexperienced	4.109	0.220
4 - Inexperienced	5.019	0.297
5 - Experienced	4.423	0.388

Weather	LS Mean	Std. Err. LS Mean
1 - cold, rainy	3.963	0.382
2 - some rain, no sun	5.221	0.783
3 - cloudy	4.120	0.251
4 - partly sunny	4.571	0.415
5 - warm, sunny	5.432	0.361

[1]Intentions to visit a visitor center is based on seven point Likert scale.

across groups (Table 18). Increasing visitor intentions thus seems more dependent upon the individual guide than whether the trip occurred before or after guide training.

Significant differences in the means were also due to the weather ($p<.0455$, $F=2.47$, Table 18). This variable was subjectively rated by the study investigator on the day of the trip and was also based on the open-ended comments of the visitors on the posttest. Examination of the least square means indicate that the worst weather day rated the lowest mean intentions while the best weather day rated the highest mean intentions to visit a New River Gorge visitor center (Table 18).

Hypothesis Five: Posttest knowledge scores of visitors who had an experienced guide will be higher than posttest knowledge scores of customers who had an inexperienced guide, both before and after training.

This hypothesis examined differences in the mean postscores due to the guides, and more specifically, their experience level. Hypothesis two previously examined mean visitor postscores, but looked for differences due to the group and treatment effects. The null hypothesis tested here was: no significant differences in the posttest knowledge scores will be due to the experience level of the guides, both before and after training.

Examination of all postscores revealed that the effect due to guide was nonsignificant even though a planned contrast revealed a significant difference between experienced and inexperienced guides ($p<.0130$, $F=6.26$, Table 11 on page 55). The least square means of visitors associated with inexperienced guides were greater than the mean postscores of visitors associated with experienced guides (Table 11). The significant effect

TABLE 19

Tests for Differences in Visitors' Posttest Knowledge Scores Across Control and Treatment Groups and Guide Experience Levels

Source	df	Control Groups F value	PR > F
GUIDE	4	0.99	.4138
- Exp vs. Inexp.	1	0.62	.4320
GROUP	1	5.87	.0165
GUIDE*GROUP	4	0.87	.4863
Error	161		
Total	170		

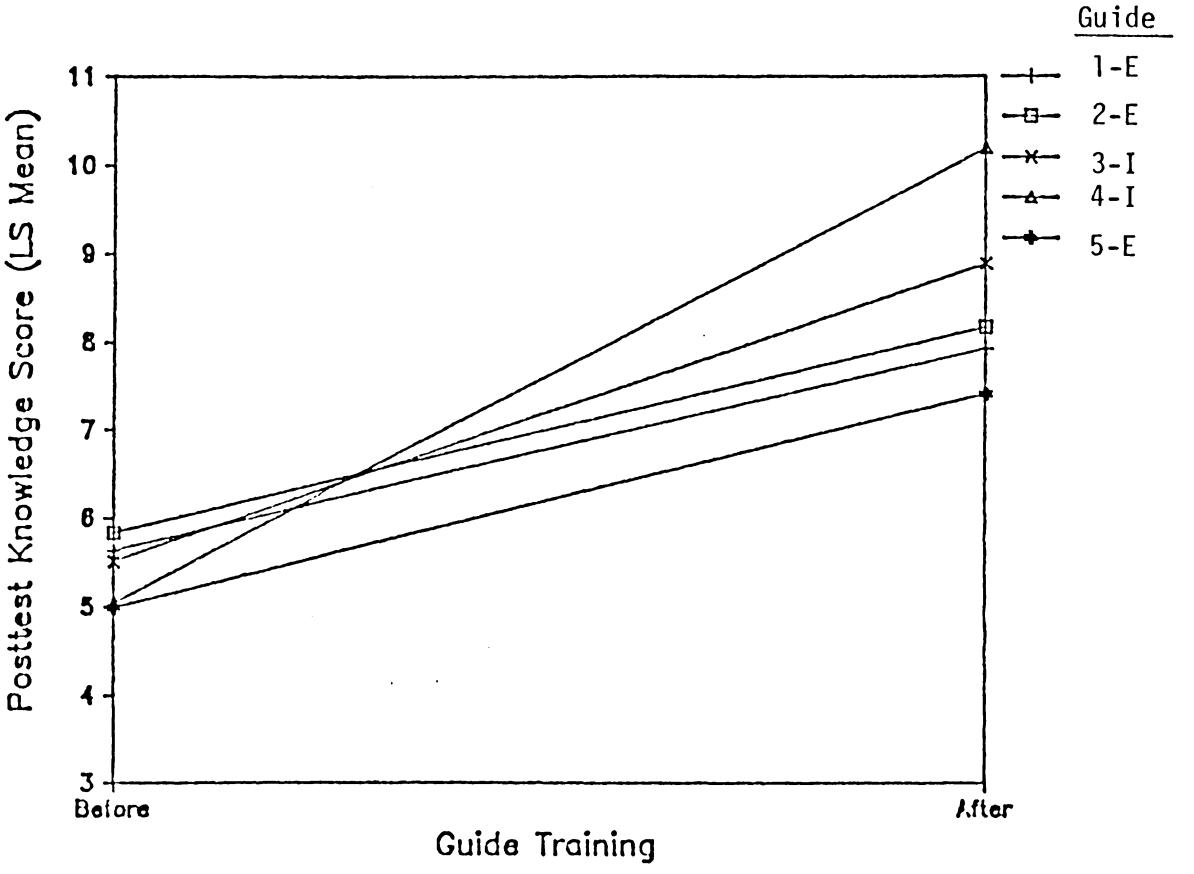
Source	df	Treatment Groups F value	PR > F
GUIDE	4	4.96	.0011
- Exp vs. Inexp.	1	18.45	.0001
GROUP	1	6.55	.0120
GUIDE*GROUP	4	3.23	.0155
Error	98		
Total	107		

Guide	All Groups LS Mean	Control Groups LS Mean	Treatment Group LS Mean
1 - Experienced	6.759	5.625	7.893
2 - Experienced	6.914	5.833	7.994
3 - Inexperienced	7.160	5.457	8.862
4 - Inexperienced	7.508	5.945	10.071
5 - Experienced	6.268	5.014	7.521

($p<.0028$, $F=2.60$) of guide by group interaction indicated that guides were not consistent across groups (Table 11). The control and treatment groups were then analyzed separately. Results indicated that significant differences in the control group least square means were not due to the guides nor their experience level (Table 19). However, significant differences in the treatment group least square means were due to the guides ($p<.0011$, $F=4.96$) and their experience level ($p<.0001$, $F=18.45$, Table 19). Treatment group least square means were greater for the customers associated with inexperienced guides than experienced guides (Table 19). The significant group effect noted here was previously discussed under Hypothesis Two. The guide by training interaction effect was also significant ($p<.0035$, $F=4.03$) and indicated that differences in the mean postscores were not consistent across guides and training (Table 12 on page 56, Figure 2).

Examination of postscore means and Figure 2 showed that before guide training, customers of inexperienced guides scored about the same as customers of experienced guides. Following guide training, the means of customers with inexperienced guides were greater than those with experienced guides (Table 12).

Hypothesis Six (a); Before training, pretest-posttest gain scores associated with experienced guides will be significantly different and greater than gain scores associated with inexperienced guides.
Hypothesis Six (b): Following training, pretest-posttest gain scores associated with experienced guides will not be significantly different from gain scores associated with inexperienced guides.



E - Experienced
I - Inexperienced

Figure 2. Mean Postscores of Visitors Associated with 5 Study Guides Before and After Guide Training.

These hypotheses examined differences in the mean pretest- posttest gain scores due to guides, and more specifically, their experience level. Hypothesis Two previously looked for differences in the mean gain scores due to the group effect. The first null hypothesis tested here was: prior to guide training, there will be no significant differences in gain score means due to the experience level of the guides. The second hypothesis, 6(b), was tested as is stated, since it is already stated in null form.

Results of the test of mean gain scores differences among control and treatment groups revealed no significant effects due to guide ($p<.2984$, $F=1.24$) or guide experience level ($p<.4756$, $F=0.51$, Table 13 on page 58). The guide by group interaction effect was nonsignificant ($p<.0994$, $F=1.99$) indicating mean gain scores of the guides were consistent across control and treatment groups (Table 13).

Results from the test for differences in the mean knowledge gain scores of the control group indicated that guides and their experience level did not explain significant differences in the means (Table 20). Knowledge gain scores of the treatment group likewise indicated that guides and their experience level did not explain significant differences in the means (Table 20). Examination of the gain score means of the control group indicated that inexperienced guides had slightly but not significantly lower least square means than the experienced guides (Table 20, Figure 3). Following training, the least square means of gain scores increased for all guides but by different amounts (Table 20). These differences were not explained by guide experience level.

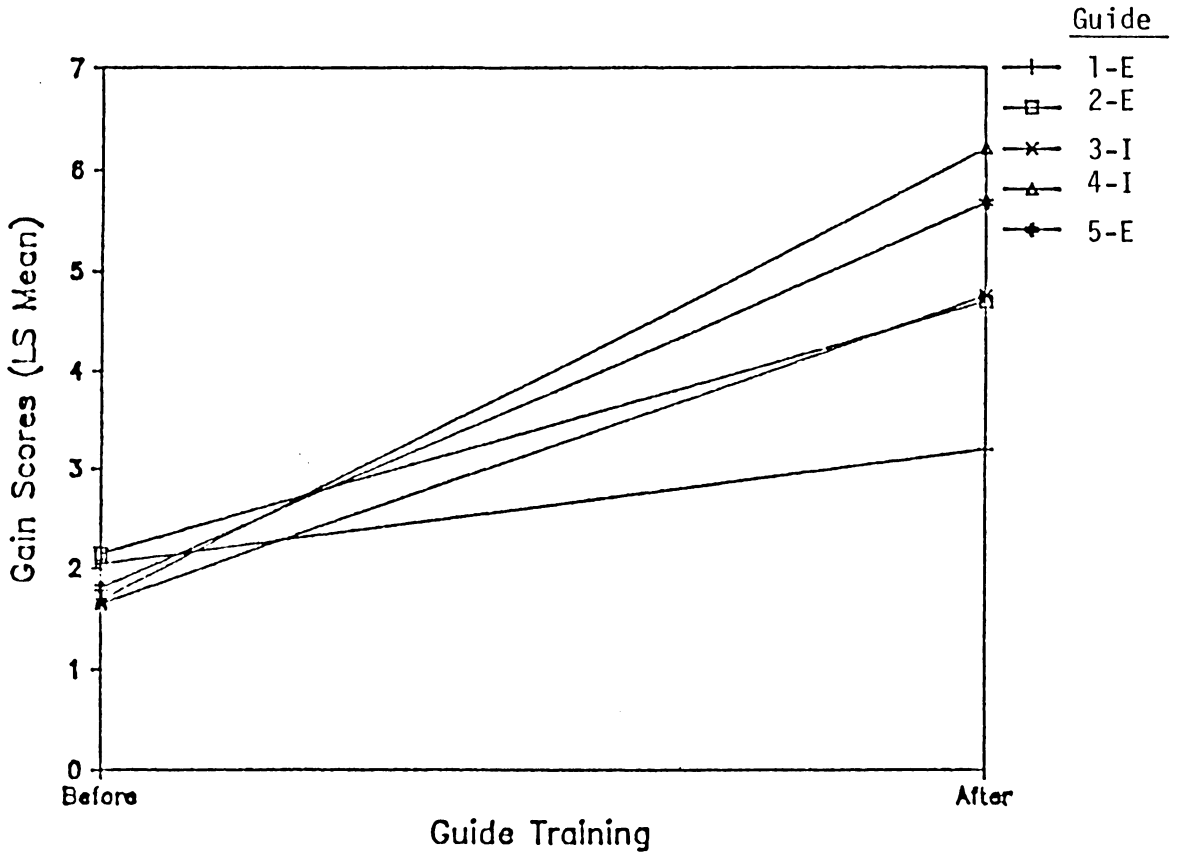
TABLE 20

Tests for Differences in Visitors' Knowledge Gain Scores
Across Control and Treatment Groups and Guide Experience Levels

Source	df	Control Groups F value	PR > F
GUIDE	4	0.17	.9507
- Exp vs. Inexp.	1	0.34	.5621
Error	79		
Total	83		

Source	df	Treatment Groups F value	PR > F
GUIDE	4	2.29	.0727
- Exp vs. Inexp.	1	2.13	.1510
Error	50		
Total	54		

Guide	All Groups LS Mean	Control Groups LS Mean	Treatment Group LS Mean
1 - Experienced	2.623	2.045	3.200
2 - Experienced	3.417	2.143	4.692
3 - Inexperienced	3.196	1.643	4.750
4 - Inexperienced	3.941	1.682	6.200
5 - Experienced	3.733	1.800	5.667



E - Experienced
I - Inexperienced

Figure 3. Mean Gain Scores of Visitors Associated with 5 Study Guides Before and After Guide Training.

DISCUSSION AND CONCLUSIONS

Summary of Findings

Acknowledgements of Potential Biases and Constraints

While several important conclusions are suggested by the findings of this study, certain potential biases and constraints should be acknowledged. For example, because field testing of the visitors was necessary, it was impossible to prevent the guides from knowing that a study was being conducted and that their customers were being evaluated. Perhaps the guides worked harder on study days to provide their customers with answers to the questionnaire in an attempt to "look good". Perhaps guides' interest in and reaction to the training program was increased by their knowledge of the visitor study. A potential "Hawthorne effect" may therefore have been operating for the guides. However, a Hawthorne effect bias would have required the guides' previous knowledge of the before and after training design and of the testing schedule, which generally was not known. A potential "Hawthorne effect" was probably not operating for the visitors. Visitors either were not aware of the study until they received the posttest or, if they did receive the pretest, were not aware that they would later be posttested. The pretest effect is accounted for by the data analysis techniques utilized.

Second, the quasi-experimental nature of the study, even with post hoc group equivalence tests, still prevents the elimination of selection bias as another potential threat to the validity of the results. The control

of many confounding variables, however, was built into the design of the study and/or accounted for by the statistical techniques used for data analysis.

Finally, the sample size and selection procedures utilized restrict generalization of results to at best the weekend visitors and guide population of Mountain River Tours during the summer months. The above potential biases and constraints were not judged by the study investigator to be present in such magnitude as to destroy confidence in the following study results.

Summary of Results

Significant differences in the amount of interpretation guides presented on the river, the amount of knowledge visitors acquired during the trip, and the visitors' overall trip rating were noted by empirical testing before and after the guide training program. Increases in the means of all three outcome variables occurred following guide training. Visitors' intentions to visit a New River Gorge Visitor Center did not significantly change following guide training.

The visitors' mean knowledge scores on a posttrip questionnaire were not found to be significantly different for experienced and inexperienced guides before guide training. Following guide training, however, the mean knowledge scores associated with inexperienced guides were significantly different and greater than knowledge scores associated with experienced guides. A significant interaction effect of guides by group was thereby noted, further indicating that increases in knowledge postscores following

training was inconsistent among guides, particularly among guides in the posttest-only treatment groups.

Before guide training, it was found that the pretest-posttest gains were slightly higher but not significantly different from the gains in knowledge of visitors associated with inexperienced guides. Following guide training, the gains in knowledge of visitors increased for all guides but to varying degrees. Differences in gain scores were not significantly explained by the guides' experience level. However, the interaction of guide experience level by group was significant, reflecting the inconsistency in gain scores associated with experienced and inexperienced guides before and after the guide training program.

Discussion

Study results confirm several, but not all, of the study hypotheses. A discussion of these results along with several additional analyses follows.

The amount of time visitors reported that guides talked about topics of the knowledge questions was measured on a single-item scale ranging from "a lot, all the time" to "never, not a word". Responses to this question indicated that guides talked about these topics more after training. In addition, visitors were asked whether they wished their guide had talked more or less about New River Gorge natural and cultural history. Examination of the frequencies of response to this question indicated that before guide training, roughly 51% of the visitors thought the amount of talk was "about right". Whereas, following training, about

64% of the visitors thought the amount of talk was "about right" (Table 8 on page 48). The percentage of visitors wishing their guide had talked more about the natural and cultural history of the New River Gorge decreased from 45% to 32% following training. The expressed desire of visitors for more talk about natural and cultural history by the guide evidently was partially satisfied following guide training, but not completely.

The pretest scores indicated an initial mean visitor knowledge level of about 31%, or 4 correct responses out of the thirteen questions before the trip and guide training. The subsequent increases in knowledge postscores were partially due to the significant pretest effect. Therefore, realistic visitor knowledge increases due to the treatment are probably estimated most accurately by the posttest-only groups. Posttest-only control group knowledge levels averaged about 38%, or 5 correct responses. This indicates an increase in knowledge levels after a "normal trip" and before guide training of about one additional correct response. Posttest-only treatment group knowledge levels averaged about 61%, or 8 correct responses out of the thirteen questions following the trip and guide training. The difference then in posttest-only visitor knowledge levels before and after guide training was about 23%, an increase of about 3 correct responses following guide training.

Another way of examining training effects on visitor knowledge is to examine the knowledge gain scores of individuals. Similar to the postscore analysis, the results indicated an increase in gain scores following training of about 23%, or 3 correct responses per individual. It should be noted that only immediate and not long term knowledge

retention was addressed by this study. Studies of long term knowledge retention are left to further research.

One of the initial objectives of the study was to determine if increased interpretation leads to increased visitor enjoyment. In fact there were significant increases from control to treatment trips in the amount of interpretation presented on the river and visitors' overall trip ratings. These outcomes were significantly correlated ($p<.0001$, Spearman's Correlation Coefficient $=.304$) as well. Responses to two of the Likert scaled questions on the posttest also addressed this objective. First visitors were directly asked if listening to and learning about the sorts of things mentioned in the 13 knowledge questions added to or detracted from the enjoyment of their trip. Frequency distributions indicated that before guide training, 58% of the visitors said it added to the enjoyment of the trip. After training this number increased to more than 70% of the visitors. Second, visitors were asked to what extent their guide's discussion of natural and cultural history contributed to the overall rating of the trip. The ratings ranged from 1="not at all" to 7="very much". The frequency of visitors choosing 6 or 7, the two response categories at the "very much" end of the scale, increased from 45% before training to 65% after training. The increase in the rating of their trip was significantly correlated ($p<.0001$, Spearman's Correlation Coefficient $= .333$) with the extent to which guides discussed natural and cultural history. This further suggests that the increase in the rating and enjoyment of the trip was in part the result of the increase in amount of interpretation following guide training.

A change in customer intentions to visit a New River Gorge visitor center was not found following training for several potential reasons.

First of all, the question addressed just one very specific behavioral intention. It is unknown to what extent guides even mentioned the visitor centers. Perhaps some guides forgot to mention this National Park Service facility. Examination of the least square means of response to the question seemed to indicate that higher means were attributed to some guides and not others, and these were consistent across experimental conditions. Some guides seemed to consistently mention the visitor centers more than others. Perhaps the existence of the visitor centers and the encouragement of visitors to stop in was not emphasized enough in the guide training program and accounts for this variability. While the effect of guides accounted for the differences in mean response to this question and the effect of training did not, perhaps future intentions can be improved with increased cooperation and interaction of the NPS and commerical guides.

It is probably easier to change a visitor's knowledge level than his behavioral intentions. It is possible that treatment group visitors simply decided that they had learned enough for one trip and did not wish to pursue it further. Control group visitors may have had an interest fostered by the trip, but the guide may not have said much. They would then be interested but perhaps unaware of what the visitor center had to offer or even possibly its location.

The significance of weather as a factor in determining visitor intentions seems to make intuitive sense. People were more likely to have intentions to travel to the visitor center on sunny, warm days than cold, rainy days. A person that has been wet and cold on the river all day probably has little intention to do anything beyond getting warm and dry.

The results found concerning the effects of guide experience upon visitors' knowledge scores should be viewed as tentative due to the small sample size comparison of two inexperienced with three experienced guides. Results, however, suggest that the inexperienced guides seemed to learn much more from the training program than did experienced guides. More precisely, following training, inexperienced guides transferred more information to the visitors than experienced guides. Experienced guides may have learned just as much at the training program, but seemed to be less likely to incorporate its presentation into their "normal trip". Perhaps inexperienced guides were more interested in learning as much about the river and its history as possible to give themselves a storyline on the river. At the same time, perhaps the experienced guides had develoepd their own stories and jokes and were less inclined to change an established behavior. Results also suggest that greater guide experience level did not produce higher knowledge scores on the control day posttests. Evidently, learning and developing white water skills does occur with an increasing number of trips down the river, but equivalent increases in understanding of the natural and cultural history of the river does not occur. Many "experienced" guides, therefore, may have developed their white water skills over the years, but are just as inexperienced with historical interpretation as the "inexperienced", first-year guides.

Implications for Management

The results of this study suggest several implications for consideration by management personnel of both the National Park Service

and commercial outfitters, such as Mountain River Tours. The benefits of a guide training program from the visitors' standpoint seem to include increased interpretation, increased knowledge, and increased trip ratings and enjoyment. Since trip ratings are always rated "very good", "excellent", or "perfect", the benefits of guide training for trip enjoyment are probably marginal. Visitor knowledge gains, however, can be highly significant, and if educating the public is a National Park Service management priority, then results of the study suggest that guide training programs can be effective at achieving these goals. Visitor knowledge of natural and cultural history as well as knowledge of the identity, services, and location of the National Park Service at the New River Gorge can be increased by guide training.

Outfitters should note the degree of interest expressed by their customers in receiving this information. Knowledgeable guides that discuss the natural and cultural history of the area generally provide more enjoyable trips. Even a marginal increase in trip ratings from "very good" to "excellent" may be the determining factor for repeat business in a highly competitive market.

This study has quantified several of the potential benefits of a guide training program, but the National Park Service must decide if they balance the costs of the inputs. Certainly outfitters should consider such staff training programs as highly desirable for all staff members, and particularly for new staff.

The increased guide/Park Service interaction at the guide training program seemed to create a more cooperative attitude and better guide understanding of the National Park Service and its purpose at the New River Gorge. Many of the guides attending the training program had little

or no previous experience or contact with Park Service employees. Familiarity and understanding seems to be a necessary development before further cooperative goals can be addressed. Perhaps this is an additional benefit that may be derived from a guide training program.

Implications for Further Research

Further research is proposed concerning several issues. The first revolves around the replication of these experimental results using a larger and more representative sample of outfitters' guides and customers, perferably with guides unaware of their participation in a study to avoid a potential Hawthorne effect. This appears necessary to increase confidence in the validity of the results, particularly for comparisons of guide characteristics such as experience level. The five guides utilized in this study may not represent the true guide population. Until replicated using more guides, results cannot be generalized to that population. There also exists the possibility that some outfitters may attract different types of customers that prefer different types of interpretive messages. Future research might therefore address the varying effectiveness of different interpretive techniques with different visitor and guide populations.

Certain objectives of this study need further refinement and testing. Specifically, the issue of guide effect on visitor behavioral intentions needs further investigation. Perhaps increasing the number of behavioral intentions addressed and tested would provide more reliable and valid results. A more complete and well developed approach seems necessary.

Guide experience level does not seem to be particularly indicative of an individual's potential success as an interpreter. Other characteristics might also be investigated. These might include various attitudes the guides hold concerning their responsibilities toward visitor education, the appropriate degree of concern with the Park Service ideals and goals, and their perceived impacts and influence over visitor behaviors.

This study was primarily concerned with short-term memory, simple recall knowledge. Further research might address the issue of long-term integrative learning and attempt to answer these questions: Do follow-up studies indicate any long-term effects from guide training on guides and visitors? How long do the benefits of training last? Do guides forget the information they receive and would continuing education seminars maintain training effects?

Another issue revolves around the maximum degree of success and training effects that might be expected from guide training. Research might investigate whether training effects are incremental in nature whereby continuous improvements might be expected with continued training. It may be, on the other hand, that training results in one large improvement with little subsequent development by the guides. Training effects may be limited by the ability of the guides or it may be limited by the length of the trip. Most trips are one day in length and perhaps there simply is not enough time during the trip to delve comfortably or effectively into subject matters very deeply. Perhaps the storyline is limted to simple noncontroversial issues because of the recreational nature of the activity.

Research should investigate the impact of various training methods as well. This study used only one intervention technique and format; perhaps

testing a variety of methods would reveal a more effective format.

The use of a questionnaire to gather data was obtrusive to many visitors, particularly those pre- and posttested. Subsequent research might opt for a posttest-only design or some other means of data collection to reduce this problem.

A final suggestion for further research efforts is a more complete investigation of the relationship of the various dependent variables such as trip rating, amount of interpretation received, visitors' knowledge scores, and visitors' experience levels. For example, it would be useful to know what degree of increase in visitor knowledge is necessary for a significant increase in visitor enjoyment and to what extent visitor experience levels affect these factors.

LITERATURE CITED

Anastasi, A. (1968). Psychological Testing. London: The MacMillan Company.

Bange, S.P. (1984). Normative influence processes among New River Gorge Boaters. Unpublished master's thesis, Virginia Polytechnic Institute and State Unversity, Blacksburg.

Campbell, D.T. and Stanley, J.C. (1966). Experimental and quasi-experimental designs for research. Chicago: Rand McNally and Co.

Cherem, C.J. (1977). The Professional Interpretor: Agent for an Awakening Giant. Association of Interpretive Naturalists (AIN) Journal 2(1), 3-16.

Cockrell, D.E., Bange, S.P., and Roggenbuck, J.W. (1984). Normative influence through interpretive communication. Journal of Environmental Education.

Cook, T.D. and Campbell, D.T. (1979). Quasi-experimentation: Design and Analysis Issues for Field Settings. Boston: Houghton Mifflin Co.

Deane, P.A. (1979). An evaluation of three nature interpretation techniques. Unpublished master's thesis, University of New Brunswick, Frederickton.

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 and Range Experiment Station, Portland, OR.

Edwards, Y. (1976). Interpretation -- what should it be? Journal of Interpretation 1(1): 4 pp.

Fazio, J.R. (1979). Communicating with the wilderness user. Bulletin #28. University of Idaho, Moscow. 65 pp.

Fazio, J.R. (1974). A mandatory permit system and interpretation for backcountry user control in Rocky Mountain National Park. (Doctoral dissertation, Colorado State University, Fort Collins, CO, 1974).

Fishbein, M. and Ajzen, I. (1975). Belief, attitude, intention, and behavior. Reading, Mass: Addison-Wesley Publishing Company.

Fritschen, J.M. (1980). An evaluation of two methods of relating environmental interpretation to urban residents of Detroit. Unpublished master's thesis, Michigan State University, E. Lansing.

Harrison, A. 91977). Getting your story across - interpreting the river resource. In Proceedings: River Recreation Management and Research Symposium. U.S.D.A. Forest Service. General Technical Report NC-28. North Central Forest Experiment Station, St. Paul, MN. p. 125-138.

Hodgson, R.W. (1984). Goal analysis and performance objectives. p. 5-32. In Supplements to a guide to cultural and environmental interpretation in the U.S. Army Corps of Engineers. U.S. Army Engineer Waterways Experiment Station. Instruction Report R-84-1. 319 pp.

Hodgson, R.W. and Fritschen, J.M. (1984). Evaluating Interpretation, p. 55-86. In Supplements to a guide to cultural and environmental interpretation in the U.S. Army Corps of Engineers. U.S. Army Engineer Waterways Experiment Station. Instruction Report R-84-1. 319 pp.

Kiely-Brocato, K. (1980). An assessment of visitor attitudes toward resource use and management. Journal of Environmental Education. 11 (Summer):29-36.

Lange, R.R. (1980). Environmental education needs assessment and evaluation manual. Vol. 1 & 2. Prepared for Colorado Department of Education.

Lime, D.W., and Lucas, R.C. (1977). Good information improves the wilderness experience. Naturalist 28(4):18-21.

Lord, F.M. (1963). Elementary models for measuring change. In C.W. Harris (Ed.), Problems in measuring change. Madison, Wis.: University of Wisconsin Press.

Mason, E.J. and Bramble, W.J. (1978). Understanding and conducting research. New York: McGraw-Hill Book Company.

Mountain River Tours. (1985). Employee responsibilities and qualification requirements, Mimeo.

Nielson, C.N. (1981). An evaluation of public interpretation programs as an aid to resource management. Leisure Resources Management Quarterly 1(1):47-58.

Nunnally, J.C. (1967). Psychometric Theory. New York: McGraw Hill.

Oliver, S.S., Roggenbuck, F.W. and Watson, A.E. (1985). Education to reduce impacts in forest campgrounds. Journal of Forestry 83(4):234-236.

Peart, B. (1979). Definition of Interpretation. Interpretation Canada 5(2):3-6.

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. U.S. Army Engineer Waterways Experiment Station, Vicksburg, Mississippi.

Putney, A.D. and Wagar, J.A. (1973). Objectives and evaluation in interpretive planning. Journal of Environmental Education 5(1):43-44.

Reagan, R. (1985). President Reagan gives his views on wide-ranging conservation issues. National Wildlife 23(4):16-19.

Risk, P.H. (1982). Education for interpreter excellence. p. 614-629. In Interpreting the Environment. New York: John Wiley and Sons, Inc.

Roggenbuck, J.W. (1979). The field experiment: a suggested method for interpretive evaluation. Journal of Interpretation 4(1):9-11.

Roggenbuck, J.W. and Berrier, D.L. (1982). A comparison of the effectiveness of two communication strategies in dispersing wilderness campers. Journal of Leisure Research 14(1):77-89.

Roggenbuck, J.W., Hall, D.F. and Oliver, S.S. (1982). The effectiveness of interpretation in reducing depreciative behavior in campgrounds. Dept. of Forestry. VPI and SU, Blacksburg. 110pp.

Sharpe, G.W. (1976). Interpreting the Environment. New York: John Wiley and Sons, Inc.

Sharpe, G.W. (1982). Interpreting the Environment. 2nd Edition, New York: Wiley and Sons, Inc.

Sharpe, G.W. and Gensler, G.L. (1978). Interpretation as a management tool. Journal of Interpretation 3(2):3-9.

Snedecor, G.W. and Cochran, W.G. (1980). Statistical Methods. 7th Edition: Iowa State University Press.

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

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

U.S. Army Engineer Waterways Experiment Station (1984). Supplements to a guide to cultural and environmental interpretation in the U.S. Army Corps of Engineers. Instruction Report R-84-1. Vicksburg, Miss.: Environmental Laboratory U.S. Army Corps of Engineers.

Wagar, J.A. 91972). 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, OR.

Wagar, J.A. (1976). Evaluating the effectiveness of interpretation. Journal of Interpretation 1(1).

Wagar, J.A. (1978). Why interpretation? Meeting the challenge. Journal of Interpretation 3(1):6-10.

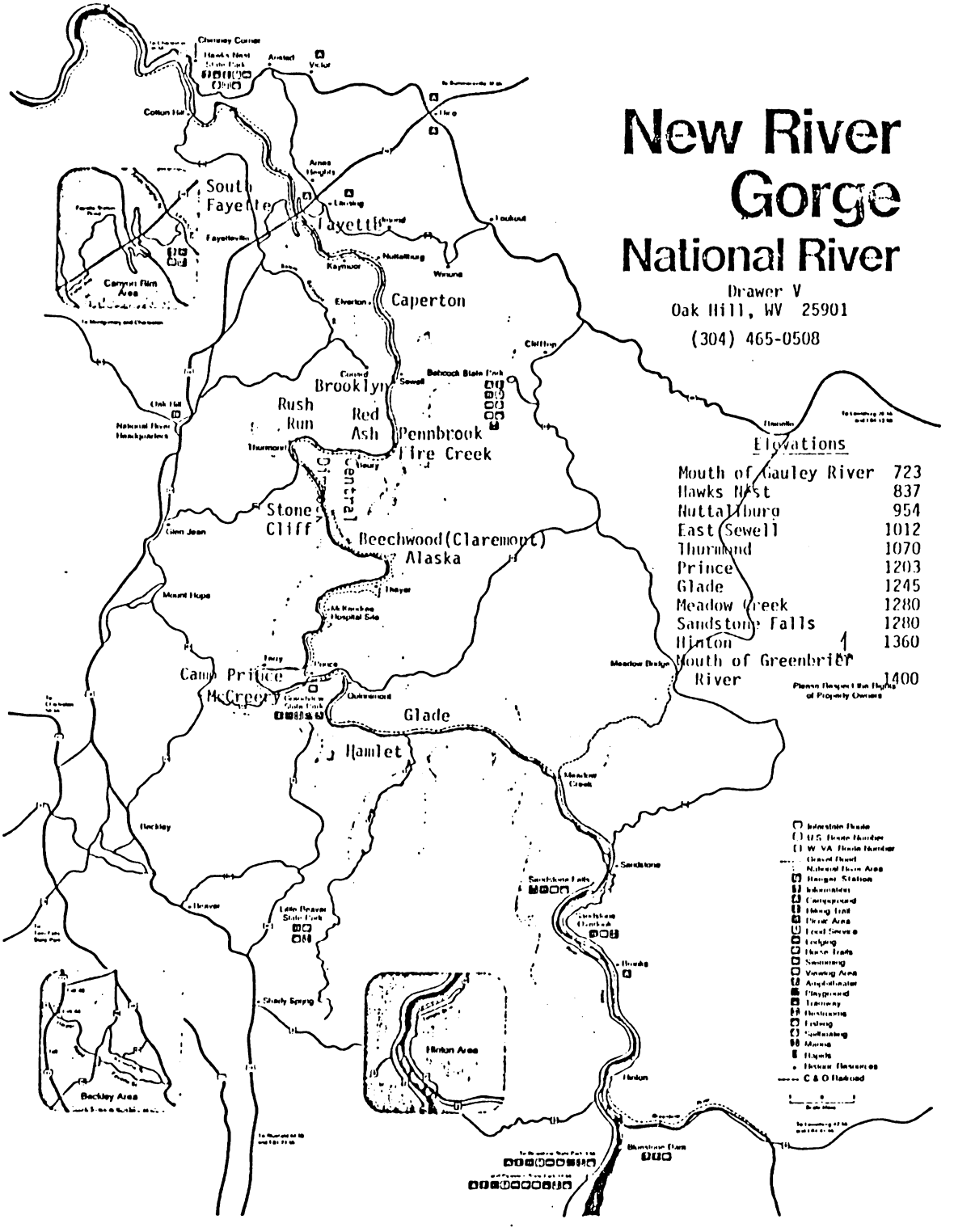
Warder, D.S. and Raile, B. (1984). Knowledge retention and satisfaction from interpretive methods. Paper presented at National Park and Recreation Association Congress, Orland, FL, Oct. 22-23.

Young, C.F. (1984). An evaluation of interpretive programs and research methods at Indiana Dunes National Lakeshore: a pilot study. Unpublished master's thesis, Michigan State University, E. Lansing.

Young, C.F. and McDonough, M.H. (1984). Interpretive Effectiveness Study: Evaluating interpretive programs and research methods at Indiana Dunes National Lakeshore, Final Report. Prepared for Eastern Parks and Monuments Association and Indiana Dunes National Lakeshore. Michigan State University, E. Lansing, MI.

Appendix A

Map of the New River Gorge National River



APPENDIX B

NEW RIVER TRIP SURVEY

To help us provide a better river trip, we need to know to what extent you are familiar with the New River Gorge. Please answer all the following questions by circling the ONE best answer.

1. The origins, or headwaters, of the New River are located in which state?
   (a) West Virginia
   (b) Virginia
   (c) North Carolina
   (d) Tennessee

2. Early explorers of the New River Gorge commented that while most rivers in eastern Virginia flowed from east to west, the New River flowed "backwards" (Northwest). This is because:

   (a) The New River follows the course of the ancient Teays River System which drained into a large inland sea.
   (b) The Gorge was carved by the last great glacial advance from Canada.
   (c) The uplifting Appalachain Mountains in West Virginia caused it to veer northward.
   (d) Ancient volcanic eruptions dammed the river and changed its course.

3. Many of the rock cliffs and huge boulders seen along the New River are this type of rock:
   (a) Granite
   (b) Sandstone
   (c) Limestone
   (d) Volcanic Gneiss

4. Most of the large boulders seen along the banks of the New River are there because they:
   (a) washed downstream.
   (b) were deposited by glaciers.
   (c) were moved by the mining activities of man.
   (d) broke off from the cliffs above.

5. The industrial boom within the New River Gorge from the 1870's to 1930's was primarily prompted by:

   (a) The completion of the Midland Trail (present Highway 60).
   (b) The opening of the Chesapeake and Ohio Railroad.
   (c) The discovery of coal.
   (d) The logging of hardwood forests.

6. Soon after the first shipment of coal in 1873, the sparsely settled New River Gorge area experienced a "boom" period. By 1900, approximately how many towns do you think existed between Thurmond and Fayette Station?
   (a) 5
   (b) 6-10
   (c) 11-20
   (d) more than 20

-over-

7. One town on the New River was not primarily a mining town, but instead a financial, entertainment, and shipping center sometimes nicknamed the "Dodge City of the East". This town was:
   (a) Sewell.
   (b) Thurmond.
   (c) Fayette Station.
   (d) Kaymoor.

8. Today there are no towns between Thurmond and Fayette Station. What was the primary cause for the decline of these mining "boom" towns?
   (a) The coal ran out.
   (b) The Depression.
   (c) Strip mining became more viable.
   (d) Lack of a competent work force.

9. Which of the following towns produced the highest revenue from freight shipment on the Chesapeake & Ohio (C&O) railroad in 1910?
   (a) Cincinnati
   (b) Charleston
   (c) Fayette Station
   (d) Thurmond

10. A mining structure that stores coal and was built over the railroad tracks for loading purposes is called the:
   (a) Tipple.
   (b) Head House.
   (c) Pump House.
   (d) Outhouse.

11. Before shipment, much of the coal mined along the New River was converted to coke, an important raw material for the steel industries to the North. How was this conversion accomplished?

   (a) The coal was burned in beehive ovens (low oxygen environment) to remove impurities.
   (b) The coal was burned in an open air (high oxygen environment) to remove impurities.
   (c) The coal was crushed and thoroughly washed with water to remove impurities.
   (d) The coal was of such quality it only required crushing and grading.

12. The New River Gorge National River was authorized in 1978 to preserve the natural and cultural resources of the National River. What government agency manages the Gorge?

   (a) West Virginia Department of Natural Resources
   (b) National Park Service
   (c) U.S. Forest Service
   (d) U.S. Army Corps of Engineers

13. Which of these facilities offers information on New River Gorge history, wildlife, scenery, and recreation opportunities in the form of slide shows, old photographs, written information, and personal services?

   (a) Canyon Rim Overlook Visitors Center
   (b) Fayetteville Chamber of Commerce
   (c) Oak Hill Public Library
   (d) West Virginia Department of Tourism

_____
Your Name or Initials

APPENDIX C

NEW RIVER TRIP SURVEY

To help us provide a better river trip, we need to know to what extent you are familiar with the New River Gorge. Please answer all the following questions by circling the ONE best answer.

1. The origins, or headwaters, of the New River are located in which state?
   (a) West Virginia
   (b) Virginia
   (c) North Carolina
   (d) Tennessee

2. Early explorers of the New River Gorge commented that while most rivers in eastern Virginia flowed from east to west, the New River flowed "backwards" (Northwest). This is because:

   (a) The New River follows the course of the ancient Teays River System which drained into a large inland sea.
   (b) The Gorge was carved by the last great glacial advance from Canada.
   (c) The uplifting Appalachain Mountains in West Virginia caused it to veer northward.
   (d) Ancient volcanic eruptions dammed the river and changed its course.

3. Many of the rock cliffs and huge boulders seen along the New River are this type of rock:
   (a) Granite
   (b) Sandstone
   (c) Limestone
   (d) Volcanic Gneiss

4. Most of the large boulders seen along the banks of the New River are there because they:
   (a) washed downstream.
   (b) were deposited by glaciers.
   (c) were moved by the mining activities of man.
   (d) broke off from the cliffs above.

5. The industrial boom within the New River Gorge from the 1870's to 1930's was primarily prompted by:

   (a) The completion of the Midland Trail (present Highway 60).
   (b) The opening of the Chesapeake and Ohio Railroad.
   (c) The discovery of coal.
   (d) The logging of hardwood forests.

6. Soon after the first shipment of coal in 1873, the sparsely settled New River Gorge area experienced a "boom" period. By 1900, approximately how many towns do you think existed between Thurmond and Fayette Station?
   (a) 5
   (b) 6-10
   (c) 11-20
   (d) more than 20

-over-

7. One town on the New River was not primarily a mining town, but instead a financial, entertainment, and shipping center sometimes nicknamed the "Dodge City of the East". This town was:
   (a) Sewell.
   (b) Thurmond.
   (c) Fayette Station.
   (d) Kaymoor.

8. Today there are no towns between Thurmond and Fayette Station. What was the primary cause for the decline of these mining "boom" towns?
   (a) The coal ran out.
   (b) The Depression.
   (c) Strip mining became more viable.
   (d) Lack of a competent work force.

9. Which of the following towns produced the highest revenue from freight shipment on the Chesapeake & Ohio (C&O) railroad in 1910?
   (a) Cincinnati
   (b) Charleston
   (c) Fayette Station
   (d) Thurmond

10. A mining structure that stores coal and was built over the railroad tracks for loading purposes is called the:
   (a) Tipple.
   (b) Head House.
   (c) Pump House.
   (d) Outhouse.

11. Before shipment, much of the coal mined along the New River was converted to coke, an important raw material for the steel industries to the North. How was this conversion accomplished?

   (a) The coal was burned in beehive ovens (low oxygen environment) to remove impurities.
   (b) The coal was burned in an open air (high oxygen environment) to remove impurities.
   (c) The coal was crushed and thoroughly washed with water to remove impurities.
   (d) The coal was of such quality it only required crushing and grading.

12. The New River Gorge National River was authorized in 1978 to preserve the natural and cultural resources of the National River. What government agency manages the Gorge?

   (a) West Virginia Department of Natural Resources
   (b) National Park Service
   (c) U.S. Forest Service
   (d) U.S. Army Corps of Engineers

13. Which of these facilities offers information on New River Gorge history, wildlife, scenery, and recreation opportunities in the form of slide shows, old photographs, written information, and personal services?

   (a) Canyon Rim Overlook Visitors Center
   (b) Fayetteville Chamber of Commerce
   (c) Oak Hill Public Library
   (d) West Virginia Department of Tourism

_____
Your Name or Initials

On a scale from 1 to 7, circle the number which best describes the way you feel.

1. How much would you say your guide talked about the topics mentioned in the previous questions?

1	2	3	4	5	6	7
Never, not a word			Sometimes			A lot, all the time

2. Does listening to and learning about these sorts of things add to or detract from the enjoyment of your trip?

1	2	3	4	5	6	7
Detract from			Neutral			Add to

3. Overall, how would you rate your trip?

1	2	3	4	5	6	7
Unacceptable	Poor	Fair	Good	Very Good	Excellent	Perfect

4. To what extent did your guides' discussion of natural and cultural history contribute to your overall rating of the trip?

1	2	3	4	5	6	7
Not at all						Very Much

5. Do you wish he had talked about New River Gorge natural and cultural history more or less?

1	2	3	4	5	6	7
Less			About right			More

6. How important is it to you that historical sites and objects along the river be preserved?

1	2	3	4	5	6	7
Not at all			Neutral			Very Important

7. Do you intend to visit any of the New River Gorge visitor centers?

1	2	3	4	5	6	7
No!			Maybe			Yes!

8. How many times have you rafted the New River?
   (a) This was my first trip
   (b) 2nd time
   (c) 3rd time
   (d) 4 or more times

-over-

10. What did you like about your trip?

11. What did you dislike about your trip?

Your Name or Initials _____ Your Guide's Name _____

Thank you for your help!

APPENDIX D

ADMINISTRATION PROTOCOL OF PRETEST QUESTIONNAIRE

The bus guide (study investigator) gave a questionnaire and pencil to each customer as he/she entered the bus. He then gave the standard outfitter pretrip speech introducing himself and the bus driver, welcoming customers to the New River, explaining river level and what to expect, explaining rules concerning no drugs or alcohol on the trip, and then introduced the questionnaire with this statement:

"One more thing before we go. We at Mountain River Tours are really interested in determining what it is our customers know about the New River Gorge. We'd like to improve our trips and we don't need to be telling you things you already know. So I'd like to ask each of you to please fill out the short survey we've handed out as best you can. I don't expect everyone to know all the answers, and that's OK. Some of these things you may discover today during our trip, other things you may not, if you're still curious at the end of the day just ask me.

You should have a survey and a pencil. Please fill it out as best you can. If you're not sure about an answer just pick whichever one you think is best. Everything is confidential, you're not being graded, you cannot flunk - so no cheating allowed! Please remember to initial your survey. We really appreciate your help!"

APPENDIX E

ADMINISTRATION PROTOCOL OF POSTTEST QUESTIONNAIRE

The bus guide (study investigator) gave a questionnaire and pencil to each customer as he/she entered the bus. To customers who had been pretested, he said:

"We're going to do this one more time, same thing with a few different questions. The first page is the same set of questions you answered this morning. Again, just answer all the questions as best you can. The second page asks things like what did you think of the trip and what could we do to improve it. Please remember to initial your survey and write down your guide's name. Your cooperation will definitely help us improve our trips."

To customers who had not been pretested the bus guide said:

"One more thing before we go. We at Mountain River Tours are really interested in finding out what it is our customers know about the New River, what you thought about the trip, and what you think we could do to improve it. So we'd really appreciate it if you'd fill out this survey answering all the questions as best you can. If you're not sure about an answer, just pick whichever one you think is best. Everything is confidential, you're not being graded, you cannot flunk - so no cheating allowed! Please remember to initial your survey and write down your guide's name. Your cooperation will definitely help us improve our trips and we really appreciate your help!

The vita has been removed from the scanned document