

# Identifying best practices for live interpretive programs in the United States National Park Service

Project Report

*for*

U.S. National Park Service

*Prepared by*

Dr. Marc J. Stern

Department of Forest Resources and Environmental Conservation  
Virginia Tech  
Phone: 540-231-7418  
Email: [mjstern@vt.edu](mailto:mjstern@vt.edu)

Dr. Robert Powell

Department of Parks, Recreation & Tourism Management  
School of Agriculture, Forest, and Environmental Sciences  
Clemson University  
Phone: 864-656-0787  
Email: [rbp@clemson.edu](mailto:rbp@clemson.edu)

Emily Martin

Department of Parks, Recreation & Tourism Management  
Clemson University

Kevin McLean

Department of Forest Resources and Environmental Conservation  
Virginia Tech

August 6, 2012

Funded by a grant from National Park Conservation Association's Center for Park Management

## **Acknowledgements**

Many people deserve our gratitude and thanks for making this study possible. We thank the National Park Service's National Education Council and its Standards Working Group as well as the NPS Comptroller's Office for their cooperation during the research. Jennifer Thomsen and Bethany Mutchler were also a part of the field research team. We would also like to thank NPCA's Center for Park Management for their financial support and further advice and comments. Finally, we would like to thank the interpretation and education staff of the National Park Service for their cooperation and enthusiasm in supporting this research project.

## **Executive Summary**

We undertook a study of live interpretive programs in the National Park Service (NPS) to examine which interpretive practices are most consistently related with desired outcomes for program attendees. A team of four researchers visited 24 units of the NPS from June through August, 2011 and observed a total of 376 live interpretive programs. In addition to conducting short interviews with interpreters immediately prior to each program to collect basic background information and to determine the interpreters' intended outcomes for their programs, the researchers monitored and recorded 56 characteristics associated with the interpreters' style of program delivery and the qualities of each program. Immediately following each program, the research team administered surveys to audience members above the age of fifteen to solicit their satisfaction with the program and their perceptions of the impacts it had on their knowledge, attitudes, behaviors, and overall park experience. Table ES-1 displays the number of programs observed and surveys collected at each park unit in the study. These units were selected to represent the wide diversity of park types, locations, and interpretive programs throughout the NPS. In total, 3,603 surveys were collected at 376 live interpretive programs.

The data collected from the monitoring of programs and the visitor surveys were then analyzed to explore relationships between the observed program characteristics and visitor-reported outcomes. Through examining a wide variety of programs across the country, the study isolated those practices that were most consistently related to desired outcomes and tested long-held assumptions about what leads to higher quality interpretive programs.

**Table ES-1. Park units included in the study and numbers of programs observed.**

<b>Park Unit</b>	<b>Resource Emphasis</b>	<b>Annual Recreation Visits</b>	<b>Programs observed</b>	<b>Visitor surveys</b>
Aztec Ruins National Monument	Cultural	37,437	2	4
Badlands National Park	Natural	977,778	19	157
Bryce Canyon National Park	Natural	1,285,492	12	133
Chaco Culture National Historical Park	Cultural	34,226	8	85
Ford's Theater National Historic Site	Cultural	662,298	20	519
Fort McHenry National Monument and Historic Shrine	Cultural	611,582	14	133
Gettysburg National Military Park	Cultural	1,031,554	21	206
Grand Canyon National Park	Natural	4,388,386	30	384
Great Smoky Mountains National Park	Mix	9,463,538	14	96
Harpers Ferry National Historical Park	Cultural	268,822	15	100
Independence National Historical Park	Cultural	3,751,007	22	156
Jefferson National Expansion Memorial	Cultural	2,436,110	16	146
Jewel Cave National Monument	Natural	103,462	20	190
Lincoln Home National Historic Site	Cultural	354,125	14	89
Manassas National Battlefield Park	Cultural	612,490	17	88
Mesa Verde National Park	Mix	559,712	14	301
Mount Rushmore National Memorial	Cultural	2,331,237	19	171
National Mall	Cultural	1,363,389	22	65
Navajo National Monument	Mix	90,696	3	23

Point Reyes National Seashore	Natural	2,067,271	9	34
San Francisco Maritime National Historical Park	Cultural	4,130,970	16	69
Ulysses S. Grant National Historic Site	Cultural	39,967	9	40
Wind Cave National Park	Natural	577,141	18	215
Yosemite National Park	Natural	3,901,408	22	199

**Program outcomes**

We used confirmatory factor analysis to aid in the development of three primary outcomes of interest for program attendees. *Satisfaction* was measured with a single survey item. *Visitor experience and appreciation* and *Behavioral change* were comprised of the average of multiple survey items, each equally weighted. These outcomes and their associated survey items are shown in Table ES-2.

**Table ES-2. Outcomes measured in the study.**

<p><b>Program outcome: Satisfaction</b>  <i>Single survey item:</i>  On a scale of 0 to 10, ten being the best, please rate your overall level of satisfaction with the program you just attend.</p>
<p><b>Program outcome: Visitor Experience and Appreciation</b>  <i>An index comprised of the average of scores on the following survey items (each measured on a 1 to 5 scale: 1 = Not at all; 5 = A great deal)</i>  To what extent did the program you just attended influence any of the following for you?</p> <ul style="list-style-type: none"> <li>• Made my visit to this park more enjoyable</li> <li>• Made my visit to this park more meaningful</li> <li>• Enhanced by appreciation for this park</li> <li>• Increased my knowledge about the program’s topic</li> <li>• Enhanced my appreciation for the National Park Service</li> </ul>
<p><b>Program outcome: Behavioral Change</b>  <i>An index comprised of the average of scores on the following survey items (each measured on a 1 to 5 scale: 1 = Not at all; 5 = A great deal)</i>  To what extent did the program you just attended influence any of the following for you?</p> <ul style="list-style-type: none"> <li>• Changed the way I will behave while I’m in this park</li> <li>• Changed the way I will behave after I leave this park</li> </ul>

Program attendees were typically highly satisfied with the programs they had just attended, averaging just below a 9.0 on the 1 to 10 scale. Scores on the *visitor experience and appreciation* index averaged a 4.43 out of 5, and the *behavioral change* index averaged a 2.96 out of 5. While behavioral intentions were less commonly influenced by programs (about 40% of respondents reported changes in their behavioral intentions after programs), the enhancement of the visitor experience was apparent throughout the sample. As such, the program and interpreter characteristics analyzed within this report do not separate good programs from bad programs. Rather, they isolate the program characteristics that are associated with great programs.

**Identified best practices**

The research suggests a suite of practices that appear to be consistently related to more positive program outcomes across contexts, resulting in the following list of best practices:

- |  |   |
|--|---|
| <ol style="list-style-type: none"> <li>1. Confidence <ul style="list-style-type: none"> <li>○ Comfort, eloquence, apparent knowledge</li> </ul> </li> <li>2. Authentic emotion and charisma <ul style="list-style-type: none"> <li>○ Passion, sincerity, charisma</li> </ul> </li> <li>3. Appropriateness for audience</li> <li>4. Organization <ul style="list-style-type: none"> <li>○ Quality of introduction, appropriate sequence, effective transitions, holistic story, clear theme, link between introduction and conclusion</li> </ul> </li> <li>5. Connection <ul style="list-style-type: none"> <li>○ Links to intangibles and universal concepts, cognitive engagement, relevance to audience, affective messaging, provocation</li> </ul> </li> </ol> | <ol style="list-style-type: none"> <li>6. Consistency</li> <li>7. Clear message</li> <li>8. Responsiveness</li> <li>9. Audibility</li> <li>10. Appropriate logistics</li> <li>11. Verbal engagement</li> <li>12. Multisensory engagement</li> <li>13. Appropriate pace</li> <li>14. Avoid focusing on knowledge gain as the program's central goal and communicating solely factual information</li> <li>15. Avoid making uncertain assumptions about the audience</li> </ol> |
|--|---|

**Management implications and recommendations**

1. NPS interpretive programs tend to produce high levels of satisfaction and experience enhancement for attendees regardless of their specific characteristics.
2. This study identified a set of practices that appear to commonly make the difference between good programs and excellent programs (with regard to visitor outcomes). We recommend these practices be incorporated into current interpretive training in the NPS where appropriate.
3. The results of our study provide a holistic perspective on practices related to positive outcomes for program attendees. However, each of these practices employed in isolation does not guarantee a high quality program. Our study instead supports the notion that interpretation is a complex phenomenon that requires competence in a range of techniques and approaches that should be responsive to different audiences and contexts. As such, we recommend maintaining the freedom for interpreters to be creative and innovative in their presentations.
4. It required considerable and iterative training, feedback, and adjustment for our team to produce consistent and reliable monitoring results. Our experiences suggest that any efforts to monitor program quality across parks will require a highly trained team with consistent membership that is external to the parks.
5. Because nearly all programs produce positive results and these data have the highest potential to be measured consistently, we recommend monitoring numbers of programs and attendees, as well as the proportion of scheduled programs that actually take place. These appear to be the most reliable measures of interpretive program health across parks. Additional measures that address audiences reached and messages communicated may also be warranted. Unless resources are made available to develop an external auditing team within the agency, we do not recommend using monitoring of these best practices to compare unit to unit.
6. We recommend that the best practices uncovered in this research be used to inform interpretive training within the NPS and by interpretive staff within parks to support reflective learning and adaptive management (iterative improvement) at the park level.

# Identifying best practices for live interpretive programs in the United States National Park Service

## Introduction

The National Park Service (NPS) describes interpretation as a communication process that is “a catalyst in creating opportunities for the audience to form their own intellectual and emotional connections with the meanings and significance inherent in the resource” (Eppley, 2012). The National Association for Interpretation has adopted a similar definition: “a mission-based communication process that forges emotional and intellectual connections between the interests of the audience and the meanings inherent in the resource” (NAI, 2012). Live interpretive programs at NPS units can serve multiple purposes. These include enhancing visitors’ enjoyment and overall experiences (Moscardo, 1999), increasing visitors’ knowledge and understanding of park resources (Ham, 1992; Tilden, 1957), fostering a sense of appreciation or other attitudes toward parks and their resources (Powell et al., 2009), and fostering both on-site and long-term stewardship behaviors (Ham, 2009).

Despite a large body of both theory and empirical research in the interpretation field, researchers have identified a lack of empirical evidence explicitly linking consensus-based best practices to visitor outcomes. These hypothesized best practices comprise those regularly suggested in common text books and other training sources alike (e.g., Ham 1992; IDP 2008; Tilden 1957; -Ward and Wilkinson 2006). In a recent review of 70 peer-reviewed research studies conducted over the last decade, Powell, Skibins, and Stern (2010) found that most evaluations of live interpretation assessed the outcomes of a single program (see also Skibins, et. al, 2012). While findings of positive outcomes across multiple studies suggest the broad efficacy of interpretation in general, no study has yet isolated the impacts of different interpretive practices and approaches upon visitor outcomes.

This study aims to close this gap in the literature through a comparative study of live interpretive programs across the NPS. In doing so, we address the following research questions:

1. What are the most common practices employed within National Park Service live interpretive programs?
2. Which practices and approaches most consistently lead to more positive outcomes for visitors?

The specific outcomes considered in this study include visitor satisfaction with programs they’ve attended, the enhancement of their experience and appreciation of the park unit and its resources, and self-reported changes in behaviors resulting from program attendance.

## Study Overview

A team of four researchers visited 24 units of the NPS from June through August, 2011 and observed a total of 376 live interpretive programs. In addition to conducting short interviews with interpreters immediately prior to each program to collect basic background information and to determine the interpreters’ intended outcomes for their programs, the researchers monitored and recorded 56 characteristics of each program associated with the interpreters’ style of program delivery and the qualities of each program. These characteristics were primarily drawn from an extensive literature review aimed at identifying published best practices in the field (Powell et al. 2010; Skibins et al., 2012)

as well as items that emerged as potentially important in pilot tests. Immediately following each program, the research team administered surveys to audience members above the age of fifteen to solicit their reactions to the program and their self-reported perceptions of the impacts the program had on their knowledge, attitudes, intentions, and overall park experience.

These data were analyzed to statistically explore relationships between the observed program characteristics and visitor-reported outcomes. Through examining a wide variety of programs across the country, the study aimed to empirically isolate which practices most consistently influenced desired outcomes, thus testing some of the field's most common assumptions about what leads to higher quality interpretive programs.

## **Methods**

### ***Site selection***

A list of relevant park criteria was created to guide which park units would be included in the research. These factors included annual visitation numbers, park location (region of the country and distance from population centers), programming focus, number of programs offered to the public, and willingness to participate in the study. In order to ensure adequate visitor attendance at interpretive programs, we only considered parks that received at least 30,000 annual recreation visits. Parks were categorized as urban, urban-proximate, or remote based on their proximity to metropolitan centers. Metropolitan areas were defined as having an urban core of at least 50,000 residents. Urban parks were located within the limits of these metropolitan areas. "Urban-proximate" parks were located outside these cores, but within a 60 mile radius of these areas. As such, they were typically in rural or suburban areas. Remote parks were located at least 60 miles from any metropolitan area. The primary interpretive themes used in each park were also considered and placed into one of three categories: predominantly cultural, predominantly natural, or a mix of the two. We aimed to maximize diversity across each of these factors. Finally, we aimed to observe at least ten programs in each park (or within nearby clusters of parks in cases such as Aztec Ruins and Navajo National Monuments) in five days or less, so we sought park units with enough programs such that this would be feasible. Twenty four park units were selected for inclusion in the study (Table 1).

We observed programs in 14 predominantly culturally-focused park units, seven predominantly nature-focused park units, and three park units with a mixed focus. This roughly mirrors the distribution of these different types of park units throughout the NPS, where roughly 30% of park units are predominantly nature-focused and roughly 60% are predominantly culturally-focused.<sup>1</sup> We visited eleven remote park units, five urban-proximate parks, and eight urban park units. This variability provides a reasonable sample from which to make generalizations to the broader population of live interpretive programs across the NPS. Park units were organized for logistical purposes by geographic region into six clusters. Teams of two researchers collected data from each park unit. One team of researchers sampled Great Smoky Mountains National Park and the mid-Atlantic, Washington D.C., and California units. The other team sampled the Southwest, Midwest, and South Dakota units.

---

<sup>1</sup> Based on a review of web pages of all park units at the time of the research ([www.nps.gov](http://www.nps.gov)).

**Table 1. Park units included in the study.**

<b>Park Unit</b>	<b>Resource Focus</b>	<b>Park Location</b>	<b>Annual Recreation Visits<sup>a</sup></b>
Aztec Ruins National Monument	Cultural	Remote	37,437
Badlands National Park	Natural	Remote	977,778
Bryce Canyon National Park	Natural	Remote	1,285,492
Chaco Culture National Historical Park	Cultural	Remote	34,226
Ford's Theater National Historic Site	Cultural	Urban	662,298
Fort McHenry National Monument and Historic Shrine	Cultural	Urban	611,582
Gettysburg National Military Park	Cultural	Urban-Proximate	1,031,554
Grand Canyon National Park	Natural	Remote	4,388,386
Great Smoky Mountains National Park	Mix	Urban-Proximate	9,463,538
Harpers Ferry National Historical Park	Cultural	Urban-Proximate	268,822
Independence National Historical Park	Cultural	Urban	3,751,007
Jefferson National Expansion Memorial	Cultural	Urban	2,436,110
Jewel Cave National Monument	Natural	Remote	103,462
Lincoln Home National Historic Site	Cultural	Urban	354,125
Manassas National Battlefield Park	Cultural	Urban-Proximate	612,490
Mesa Verde National Park	Mix	Remote	559,712
Mount Rushmore National Memorial	Cultural	Remote	2,331,237
National Mall	Cultural	Urban	1,363,389
Navajo National Monument	Mix	Remote	90,696
Point Reyes National Seashore	Natural	Urban-Proximate	2,067,271
San Francisco Maritime National Historical Park	Cultural	Urban	4,130,970
Ulysses S. Grant National Historic Site	Cultural	Urban	39,967
Wind Cave National Park	Natural	Remote	577,141
Yosemite National Park	Natural	Remote	3,901,408

<sup>a</sup> Annual visitation from 2010 (<http://www.nature.nps.gov/stats/>)

### ***Sampling and data collection***

Individual live interpretive programs served as the unit of analysis for this study. Programs were selected first and foremost based on their time and location within each park unit to maximize the number of programs observed at each park unit. We also aimed to maximize variability in the types of programs (both with regard to subject matter -- natural vs. cultural -- and types of delivery -- guided walks vs. campfire programs vs. hands-on activities, etc.). Regular programs were selected over children's

programs whenever an explicit alternative was available, as adult respondents were the targets of visitor surveys.

Throughout the research, the same procedure was followed for observing all programs. Upon arrival at the program site, a brief interview was conducted with the interpreter. Interview questions included a battery of intended programmatic outcomes, questions about program development, and others about the preparation and the level of enthusiasm of the interpreter. The interviews also collected basic background information about the interpreter, which included age, gender, and interpretation experience. These interviews were conducted on all but 15 programs. In those cases, time did not allow for the interviews to take place. Basic information about the program itself was recorded by the observer, including time, location, type, topic focus, and size and age breakdown of audience.

We asked each interpreter to make a brief announcement at the beginning of each program so that visitors were aware of our presence. This did not happen consistently, however. When it did not, an announcement was made either by the interpreter or the researcher at the end of the program. Visitors above the age of 15 were asked to remain after the program to complete a short survey. In programs that were particularly large (more than 50 attendees), the researchers employed systematic sampling whenever possible – for example, selecting every  $n^{\text{th}}$  row to complete surveys at Ford’s Theatre. In these cases the researchers chose the sample interval in attempt to target at least 20 respondents.

Throughout the program, researchers maintained an unobtrusive presence within the group, acting simply as another member of the audience. Observation sheets were completed by the researchers immediately following each program.

Throughout the duration of all field work, researchers would occasionally attend programs together to ensure reliability and consistency in scoring each variable. Occasional check-ins were also completed between team members to ensure that observation techniques were consistent, to clarify questions about scoring certain variables, and to add variables that were deemed relevant to the research. No new variables were added after the first week of fieldwork.

We collected 3603 surveys from visitors who attended 376 programs. Table 2 presents a summary of the programs we attempted to observe, those we actually observed, and the total number of surveys collected at each program. The difference between the “programs attempted” and “programs observed” columns indicates the number of scheduled programs we attempted to observe that did not take place as planned.

**Table 2. Programs observed and total number of surveys collected.**

<b>Park unit</b>	<b>Programs attempted</b>	<b>Programs observed</b>	<b>Surveys collected</b>
Aztec Ruins National Monument	4	2	4
Badlands National Park	22	19	157
Bryce Canyon National Park	12	12	133
Chaco Culture National Historical Park	9	8	85
Ford's Theater National Historic Site	20	20	519
Fort McHenry National Monument and Historic Shrine	23	14	133
Gettysburg National Military Park	26	21	206
Grand Canyon National Park	30	30	384
Great Smoky Mountains National Park	19	14	96
Harpers Ferry National Historical Park	21	15	100
Independence National Historical Park	36	22	156
Jefferson National Expansion Memorial	22	16	146
Jewel Cave National Monument	20	20	190
Lincoln Home National Historic Site	18	14	89
Manassas National Battlefield Park	20	17	88
Mesa Verde National Park	14	14	301
Mount Rushmore National Memorial	23	19	171
National Mall	47	22	65
Navajo National Monument	8	3	23
Point Reyes National Seashore	12	9	34
San Francisco Maritime National Historical Park	20	16	69
Ulysses S. Grant National Historic Site	15	9	40
Wind Cave National Park	18	18	215
Yosemite National Park	29	22	199
<b>Totals</b>	<b>488</b>	<b>376</b>	<b>3603</b>

**Measurement*****Dependent variables: Outcomes***

Dependent variables represent measurements of the outcomes of interest. The outcome variables in the study were comprised of retrospective assessments provided by program attendees on surveys administered immediately following their programs.<sup>2</sup> While interpretation may produce multiple outcomes, we focused primarily on visitor satisfaction and shifts in knowledge, attitudes, and behavioral intentions relevant to the park experience.

<sup>2</sup> Our original research design also included administering shorter pre-experience surveys at different, but similar programs across the parks in our sample. These surveys contained two batteries of survey items that could be compared to the post-experience surveys to create a control group against which to compare outcomes. Unfortunately, an insufficient number of these surveys were administered at most parks to create a reliable control group. As a result, we did not include these data in further analyses.

Overall satisfaction with the program was measured on a scale from 0 to 10, with 0=Terrible and 10=Excellent. An additional battery of survey items provided response prompts for the following question: “To what degree did the program you just attended influence any of the following for you?” Response categories were comprised of a 5-point Likert-type scale, with answer choices: Not at all (1), A little (2), Somewhat (3), A moderate amount (4), and A great deal (5). The survey items included:

- Made me think deeply
- Made me reflect on my own life
- Enhanced my appreciation for this park
- Enhanced my appreciation for the National Park Service
- Made me more likely to avoid harming park resources
- Increased my knowledge about the program’s topic
- Made my visit to this park more enjoyable
- Made my visit to this park more meaningful
- Changed the way I will behave while I’m in this park
- Changed the way I will behave after I leave this park
- Made me want to tell others about what I learned
- Made me care more about this park’s resources
- Made me care more about protecting places like this

These items were developed based on key literature (e.g., Ham, 1992; Moscardo, 1957, Tilden 1957; Ward and Wilkinson 2006) and extensive input from NPS staff. This input included interviews and focus groups with the National Education Council; a focus group and associated surveys conducted with NPS interpreters at the NAI National Workshop in Las Vegas, November 2010; and two surveys conducted in 2010 and 2011 with NPS superintendents and supervisors of interpretation, respectively (see Stern and Powell 2011).

### ***Independent variables: predictors***

Independent variables refer to characteristics that are expected to exhibit statistical relationships with outcomes. In this report, we refer to them as “predictor variables.” Predictor variables within the study existed at two levels. Our primary interest is in program-level characteristics. However, the characteristics of the visitors were also collected to determine their relationships with program outcomes. Visitors reported their age, group types, whether they had attended a ranger-led program at the park before, their sex, their zip code, their ethnicity, and whether there were children in their group. Zip codes were used to determine if a visitor could be considered a park neighbor, or “local,” or not. Local visitors may or may not respond differently to programs than non-local visitors (Stern et al., 2012). Local visitors were defined as those whose zip codes were located within a 1 hour driving distance from the park’s nearest entrance, based on the center of the zip code’s mapped land area.

The primary variables of interest included both interpreter characteristics and the interpretive practices employed during a program. These characteristics were primarily drawn from an extensive literature review aimed at identifying best practices in the field (Powell et al., 2010; Skibins et al., 2012) as well as characteristics identified by interpretive experts within the NPS and ranked highly by interpretive staff in surveys (Stern and Powell, 2011) and items that emerged as potentially important in pilot tests. Interpreter characteristics focus upon the appearance, identity, and overall styles of the interpreters themselves (Table 3). Program characteristics are more specifically-related to qualities of the programs themselves (Table 4). Citations are provided where characteristics were drawn from the literature.

Examples of each of these characteristics from our qualitative observations are shared in Appendix A to further clarify how they manifested in the field.

We also collected details pertaining to the program including the experience level and demographics of the interpreter, their intended outcomes for their programs, and their level of excitement about the particular program they were about to deliver. We also tracked information on the type of program, its focus (natural vs. cultural/historical vs. both), and other unexpected circumstances that could impact program outcomes (e.g., weather). We also estimated the number of attendees at each program and the ratio of youth (ages 15 and under) to adults.

Extensive pilot testing aided in instrument development and refinement and enhanced the reliability of measurement across the research team. Prior to the field research, we observed video recorded interpretive programs from an undergraduate interpretation class at Virginia Tech. These programs were used to establish consistent measurement of each relevant characteristic. Programs were viewed repeatedly and scores were compared among team members on each characteristic. These exercises were also used to refine the scoring of several variables.

From this testing, a preliminary assessment sheet was developed. These assessment sheets were further pilot-tested collectively by the research team at Great Smoky Mountains National Park in May of 2011, where the team observed three live interpretive programs. Extensive discussion allowed us to further refine definitions and observation techniques for each of the characteristics under study. For each measure, we aimed to maximize the range of possible scores to enhance variability in the findings. However, existing definitions from the literature and results of pilot-testing limited most scales to four or fewer points. Pilot testing revealed that the middle-points on larger scales for many variables were not easily differentiated in a consistent manner by the research team. As a result, the scoring for each item varies to maximize the potential range of scores while maintaining inter-rater reliability. Binary scores were used in cases where the most appropriate measure was to indicate presence or absence.

**Table 3. Interpreter characteristics observed in the study, their definitions, and operationalization.**

<b>Interpreter characteristic</b>	<b>Definition</b>	<b>Scoring</b>
<b>Professional appearance</b>	The extent to which the interpreter appears properly dressed and groomed.	0 = Interpreter appears disheveled or unkempt and is not professionally dressed 1 = Interpreter appears well-groomed and is professionally dressed
<b>Comfort of the interpreter</b> (Lewis 2008; Moscardo, 1999; Ward and Wilkinson, 2006)	Degree to which the interpreter presenting the program seems comfortable with the audience and capable of successfully presenting the program without apparent signs of nervousness or self-doubt.	1 = Interpreter seems scared, nervous, or unable to lead the program 2 = Interpreter seems nervous and struggles with much of the program 3 = Interpreter seems comfortable, but might become uncomfortable at times 4 = Interpreter is not nervous and handles the program with ease

<p><b>Responsiveness</b> (Jacobson, 1999; Knudson et al., 2003; Lewis, 2008)</p>	<p>The extent to which the interpreter interacts with the audience, collects information about their interests and backgrounds, and responds to their specific questions and requests or non-verbal cues.</p>	<p>NA = Not able to observe (e.g., large programs in dark theatres) 1 = Interpreter is aloof or averse to the visitors' presence 2 = Interpreter is somewhat responsive to visitors' questions/body language 3 = Interpreter was very responsive to the audience</p>
<p><b>Inequity</b> (Ham and Weiler, 2002)</p>	<p>The presence of unequal attention devoted to certain attendees and not others through greater interaction or attentiveness.</p>	<p>1 = Interpreter did not pay equal attention to all audience members. 0 = No inequity issues.</p>
<p><b>Humor quality</b> (Ham and Weiler, 2002; Knapp and Yang, 2002; Regnier et al., 1992)</p>	<p>How funny is the interpreter overall? Does the audience react positively to the interpreter's use of humor and seem to enjoy it?</p>	<p>1 = Not funny at all 2 = A little funny 3 = Moderately funny 4 = Hilarious</p>
<p><b>Humor quantity</b></p>	<p>The extent to which the interpreter attempts to use humor, sarcasm, or jokes to share the topic with the visitor, regardless of their success.</p>	<p>1 = Interpreter attempts no humor throughout the presentation 2 = Interpreter rarely uses humor 3 = Interpreter uses an equal mix of humor and non-humor to convey the message 4 = Interpreter is mostly trying to be humorous 5 = Interpreter uses humor as the primary vehicle to convey their message</p>
<p><b>Sarcasm</b></p>	<p>The degree to which the interpreter used sarcasm (the use of mocking, contemptuous, or ironic language or tone) or self-deprecation that was not meant to be serious, as a part of presenting their program.</p>	<p>1 = Not at all 2 = Done to some extent 3 = A central feature of the delivery style</p>
<p><b>Charisma</b> (Ward and Wilkinson, 2006)</p>	<p>A general sense of the overall likeability/charisma of the interpreter, commonly recognized by seemingly genuine interaction with the visitors, including smiling, looking people in the eye, and having an overall appealing presence.</p>	<p>1 = Not likeable/found interpreter irritating 2 = Somewhat off-putting 3 = Neither liked or disliked interpreter 4 = More or less liked interpreter 5 = Found interpreter very likeable/charismatic</p>

<p><b>Sincerity</b> (Ham, 2009)</p>	<p>The degree to which the interpreter seems genuinely invested in the messages he or she is communicating, as opposed to reciting information, and seems sincere in the emotional connection they may exude to the message and/or the resource. In other words, the extent to which the interpretation was delivered through authentic emotive communication.</p>	<p>1 = Interpreter seemed to only be going through the motions, with no real emotional connection or sincerity  2 = Interpreter seemed somewhat connected through the words they used, though their mannerisms or intonation didn't corroborate their words.  3 = Interpreter seemed mostly sincere with authentic emotive communication for most of the program  4 = Communication was clearly sincere and authentic throughout the program, as evidenced by words, gestures, intonation, or other mannerisms</p>
<p><b>Passion</b> (Beck and Cable, 2002; Ham and Weiler, 2002; Moscardo, 1999)</p>	<p>The interpreter's apparent level of enthusiasm for the material, as opposed to a bored or apathetic attitude toward it. The overall vigor with which the material is presented.</p>	<p>1 = Interpreter seems completely detached/disinterested from the program  2 = Low levels of passions  3 = Interpreter shows moderate levels or sporadic instances of passion  4 = Pretty high levels of passion overall  5 = Interpreter seems extremely passionate about the program</p>
<p><b>Personal sharing</b> (Jacobson, 1999)</p>	<p>The degree to which the interpreter shared personal insights or experiences, answered questions about themselves for the audience, or provided their own opinion on topics or events relevant to the program.</p>	<p>1 = Interpreter did not share any personal information about themselves with the audience  2 = Interpreter shared minimal personal information or viewpoints  3 = Interpreter shared a large amount of personal information and perspective  4 = Interpreter's personal life/point of view is explicitly the central focus of the experience (used themselves as the primary framework for the program)</p>
<p><b>Apparent knowledge</b> (Ham and Weiler, 2002; Lewis, 2008; Ward and Wilkinson, 2006)</p>	<p>The degree to which the interpreter appears to know the information involved in the program, the answers to visitors questions, and has local knowledge of the area and its resources.</p>	<p>1 = Interpreter seems not at all knowledgeable (unsure of facts or has a hard time recalling the information intended for the program)  2 = Interpreter seems somewhat knowledgeable, but appears to forget a few things or leave out important details  3 = Interpreter appears more or less knowledgeable without any major hiccups or uncertainty throughout the program.  4 = Interpreter's presentation of facts and information during the program is flawless</p>

<b>Audibility</b>	The extent to which the interpreter can clearly be heard and understood by the audience.	1 = Interpreter could not be heard by the audience during the majority of the program 2 = Interpreter could be clearly heard for the majority of the program, but wasn't audible during some parts 3 = Interpreter could be clearly heard throughout the entire program
<b>Eloquence</b> (Lewis, 2008)	The extent to which the interpreter spoke clearly and articulately, and did not mumble or frequently use filler words such as "um" or "like."	1 = Interpreter stumbled on their speech throughout their entire program and was hard to understand 2 = Interpreter had some minor issues with mumbling or unclear speech 3 = Interpreter had no such issues during the program 4 = Interpreter was exceptionally eloquent
<b>Impatience</b>	Did the interpreter show any explicit impatience toward audience members?	1 = Interpreter was explicitly impatient with the audience 0 = No issues noted
<b>Formality</b>	The degree to which the interpreter was very formal and official vs. casual and laid back about the presentation.	1 = Interpreter was extremely casual 2 = More casual than formal 3 = Interpreter was neither explicitly casual nor formal 4 = More formal than casual 5 = Interpreter was entirely formal
<b>False assumption of the audience</b>	At any point during the program, did the interpreter make assumptions of the audience's attitudes or knowledge that could have easily been false?	1 = No problem with false assumptions 2 = Some minor false assumptions that likely did not detract from the quality of the program 3 = Obvious false assumptions that made the experience less enjoyable or meaningful
<b>Character acting</b>	The degree to which role playing or character acting is incorporated into the program, either to add authenticity or to help tell a story.	0 = Interpreter does no character role playing during the program, he/she is simply leading the program 1 = Interpreter acts like one or more characters during parts of the program 2 = Interpreter is in full costume or does not break character at any point during the program
<b>Primary identity</b> (Ham and Weiler, 2002; Ham, 2002, Knapp and Yang, 2002; Larsen, 2003; Wallace and Gaudry, 2005)	<b>Friend:</b> outwardly friendly, casual, approachable, mingles informally	1 = primary identity; 0 = not
	<b>Authority figure:</b> emphasizes own role as a park ranger and focuses on rules, regulations, and/or authority to communicate	1 = primary identity; 0 = not
	<b>Walking encyclopedia:</b> Focused on conveying a large volume of facts	1 = primary identity; 0 = not

<b>Questionable information</b>	Obvious factual inaccuracy (incorrect or inaccurate information) or false attribution (unfounded claims about others, e.g., “the native people were happy to hand over their land so a National Park could be formed.”)	1 = present 0 = not present
<b>Bias</b>	Did the interpreter share any apparent bias or strong opinion with potential effects on relationships with audience members?	1 = yes 0 = no

**Table 4. Program characteristics observed in the study, their definitions, and operationalization.**

<b>Program characteristic</b>	<b>Definition</b>	<b>Scoring</b>
<b>Introduction quality</b> (Brochu and Merriman, 2002; Ham, 1992; Jacobson, 1999)	Degree to which the introduction captured the audience’s attention and oriented (or pre-disposed) the audience to the program’s content and/or message.	3= Oriented audience and captured attention 2= Minimally oriented audience; did not necessarily capture attention 1= Poorly executed
<b>Appropriate logistics</b> (Jacobson, 1999; Knudson et al., 2003)	Degree to which basic audience and program needs were met (i.e., restrooms, weather, technology, accessibility, shade, etc).	4= Well planned and appropriate 3= Audience/program needs mostly addressed 2= Needs marginally addressed 1= Needs not met
<b>Comfort of the audience</b> (Brochu and Merriman, 2002; Jacobson, 1999; Milton et al., 1995, in Knudson et al., 2003)	Degree of physical comfort of the audience.	4= Audience very comfortable 3= Audience comfortable 2= Audience uncomfortable 1= Audience very uncomfortable
<b>Appropriate for audience</b> (Beck and Cable, 2002; Jacobson, 1999; Knudson et al., 2003)	Degree to which the program aligned with audience’s ages, cultures, and level of knowledge, interest, and experience.	5= Very appropriate 4= Appropriate 3= Moderately appropriate 2= Only slightly appropriate 1= Not appropriate
<b>Appropriate sequence</b> (Beck and Cable, 2002; Ham, 1992; Jacobson, 1999; Larsen, 2003)	Degree to which the program followed a logical sequence.	4= Enhanced messaging 3= Appropriate 2= Choppy 1= Detracted from messaging
<b>Transitions</b> (Beck and Cable, 2002; Brochu and Merriman, 2002; Ham, 1992; Jacobson, 1999; Larsen, 2003)	Degree to which program used appropriate transitions that kept the audience engaged and did not detract from the program’s sequence.	4= Enhanced messaging and were smooth 3= Appropriate 2= Forced or irrelevant 1= Detracted from messaging or not present

<p><b>Intangibles and universals</b> (NPS Module 101; Beck and Cable, 2002; Brochu and Merriman, 2002; Ham, 1992; Knudson, et al., 2003; Larsen, 2003; Lewis, 2005; Moscardo, 1999; Tilden, 1957; Ward and Wilkinson, 2006)</p>	<p>Communication connected tangibles to intangibles and universal concepts. Intangibles: stories, ideas, meanings, or significance that tangible resources represent Universals: concepts that most audience members may identify with</p>	<p>5= Extensively developed; powerful concepts 4= Well developed 3= Present but weak 2= Difficult to detect or slightly used 1= Clearly not present</p>
<p><b>Multisensory</b> (Beck and Cable, 2002; Knudson et al., 2003; Lewis, 2005; Moscardo, 1999; Tilden, 1957; Veverka, 1998)</p>	<p>Degree to which the program intentionally and actively engaged more than just basic sight and sound.</p>	<p>3= Explicit/purposeful inclusion of two sense beyond sight and sound 2= Actively incorporated a sense beyond passive use of sight and sound, or actively focused upon either of these senses as a vehicle for conveying the message (e.g., "close your eyes and listen") 1= Primarily a talk in which the ranger did not explicitly use multiple sense beyond passive use of sight (scenery/objects) and sound (words)</p>
<p><b>Physical engagement</b> (Beck and Cable, 2002; Knudson, et al., 2003; Lewis, 2005; Moscardo, 1999; NPS Module 101; Sharpe, 1976; Tilden, 1957)</p>	<p>Degree to which the program physically engaged audience members in a participatory experience; i.e., through touching or interacting with resource.</p>	<p>4= Central programming element 3= Occurred multiple times 2= Minimal effort to engage 1= No efforts</p>
<p><b>Verbal engagement</b> (Knudson, et al., 2003; Moscardo, 1999; Sharpe, 1976; Tilden, 1957; Veverka, 1998)</p>	<p>Degree to which the program verbally engaged audience members in a participatory experience; i.e., dialogue (a two-way discussion).</p>	<p>5= Central programming element 4= Occurred multiple times 3= Modestly engaged 2= Minimal effort to engage 1= No efforts</p>
<p><b>Cognitive engagement</b> (Knudson, et al., 2003; Moscardo, 1999; Sharpe, 1976; Tilden, 1957; Veverka, 1998)</p>	<p>Degree to which the program cognitively engaged audience members in a participatory experience beyond simply listening; i.e. calls to imagine something, reflect, etc.</p>	<p>5= Central programming element 4= Occurred multiple times 3= Modestly engaged 2= Minimal effort to engage 1= No efforts</p>
<p><b>Multiple activities</b> (Knapp and Benton, 2004; Moscardo, 1999; Ward and Wilkinson, 2006)</p>	<p>Degree to which the program consisted of a variety of activities and opportunities for direct audience involvement (not including dialogue).</p>	<p>4= 2+ primary activities included 3= 2+ secondary activities included 2= One secondary activity included 1= One activity only</p>
<p><b>Props</b> (Jacobson, 1999; Knapp and Benton, 2005; Ham, 1992)</p>	<p>A visual aide beyond a screen-based slideshow.</p>	<p>1 = Prop(s) used 0 = Not used</p>

<p><b>Relevance to audience</b> (Beck and Cable, 2002; Brochu and Merriman, 2002; Ham, 1992; Jacobson, 1999; Knapp and Benton, 2004; Lewis, 2005; Moscardo, 1999; NPS Module 101; Sharpe, 1976; Tilden, 1957; Veverka, 1998)</p>	<p>Degree to which the program explicitly communicated the relevance of the subject to the lives of the audience.</p>	<p>5= Major focus of messaging 4= Well developed efforts 3= Moderate efforts 2= Minimal efforts 1= No efforts</p>
<p><b>Affective messaging</b> (Jacobson, 1999; Lewis, 2005; Madin and Fenton, 2004; Tilden, 1957; Ward and Wilkinson, 2006)</p>	<p>Degree to which the program communicated emotion (in terms of quantity, not quality).</p>	<p>5= Central programming element 4= Frequent and repeated messages 3= Occasional messages 2= Minimal effort to include messages 1= Messages absent</p>
<p><b>Fact-based messaging</b> (Frauman and Norman, 2003; Jacobson, 1999; Lewis, 2005; Tilden, 1957; Ward and Wilkinson, 2006)</p>	<p>Degree to which the program communicated factual information.</p>	<p>1 = Messaging was solely fact-based 0 = Messaging was not solely fact-based (incorporated affective messaging)</p>
<p><b>Surprise</b> (Beck and Cable, 2002; Moscardo, 1999)</p>	<p>Degree to which the program used the element of surprise in communication. This could include “aha” moments or unexpected or contrasting messages.</p>	<p>3= Major element 2= Minor element 1= Not used</p>
<p><b>Novelty</b> (Beck and Cable, 2002; Frauman and Norman, 2003; Knapp and Benton, 2004; Moscardo, 1999)</p>	<p>Degree to which the program presented novel ideas, techniques, or viewpoints as an element of communication; i.e., using a device not usually associated with or related to resource.</p>	<p>3= Major element 2= Minor element 1= Not used</p>
<p><b>Provocation</b> (Beck and Cable, 2002; Brochu and Merriman, 2002; Knudson, et al., 2003; Tilden, 1957)</p>	<p>Degree to which the program explicitly provoked participants to personally reflect on content and its deeper meanings.</p>	<p>4= Powerful and explicit inclusion 3= Occasional inclusion 2= Isolated or vague inclusion 1= No attempt made</p>
<p><b>Multiple viewpoints</b> (Beck and Cable, 2002; Brochu and Merriman, 2002; Tilden, 1957)</p>	<p>Degree to which the program explicitly acknowledged multiple perspectives or uncertainty within a theme or message. (Primarily for controversial messaging; when an argument is made, was a relevant counter-argument provided?)</p>	<p>3= Multiple viewpoints developed; none given clear priority 2= Primarily one viewpoint, with some focus on others 1= No effort NA = not applicable</p>

<p><b>Holistic storytelling</b> (Beck and Cable, 2002; Larsen, 2003; Tilden, 1957)</p>	<p>Degree to which the program aimed to present a holistic story (with characters and a plot) as opposed to disconnected pieces of information.</p>	<p>5= Holistic story used throughout; all messaging tied to story 4= Holistic story present; some info did not relate to story 3= Equal mix of storytelling and factual information, no single, holistic story 2= Factual information primarily used; some stories used to create relevance. 1= Facts and information primarily; no attempt at storytelling.</p>
<p><b>Place-based messaging</b> (Beck &amp; Cable, 2002; Knudson, et al., 2003; Lewis, 2005; Moscardo, 1999; NPS Module 101; Sharpe, 1976)</p>	<p>Degree to which the program emphasized the connection between the visitor and the site/resource.</p>	<p>5= Central focus of messaging 4= Well-developed connection through repetition and engagement 3= Moderately emphasized through repetition or engagement 2= Slightly developed verbally 1= Not developed</p>
<p><b>Introduction and conclusion linkage</b> (Beck and Cable, 2002; Brochu and Merriman, 2002; Larsen, 2003)</p>	<p>Degree to which program connected conclusion back to the introduction in an organized or cohesive way (i.e., program “came full circle.”)</p>	<p>4= Intro and conclusion were linked in a cohesive way that enhanced messaging 3= Intro and conclusion were linked, but didn’t necessarily enhance messaging 2= Intro and conclusion were weakly linked 1= Intro and conclusion were disconnected from each other</p>
<p><b>Clear theme</b> (Beck and Cable, 2002; Brochu and Merriman, 2002; Ham, 1992; Jacobson, 1999; Knudson, Cable, and Beck, 2003; Larsen, 2003; Lewis, 2005; Moscardo, 1999; Sharpe, 1976; Veverka, 1998; Ward and Wilkinson, 2006)</p>	<p>Degree to which the program had a clearly communicated theme(s). A theme is defined as a single sentence (not necessarily explicitly stated) that links tangibles, intangibles, and universals to organize and develop ideas.</p>	<p>4= Theme is clearly developed and communicated 3= Easy to detect, but not well developed 2= Difficult to detect, present but at least somewhat ambiguous 1= Unclear/not present</p>
<p><b>Central message</b> (Beck and Cable, 2002; Brochu and Merriman, 2002; Jacobson, 1999)</p>	<p>Degree to which program’s message(s) was clearly communicated; i.e., the “so what?” element of the program.</p>	<p>4= Clearly communicated and well developed 3= Easy to detect, but not well developed 2= Difficult to detect, ambiguous 1= Unclear/not present</p>
<p><b>Consistency</b> (Beck and Cable, 2002; Ham, 1992)</p>	<p>Degree to which the program’s tone and quality were consistent throughout the program</p>	<p>3=Consistent 2=Some shift in either tone or quality during the program 1= Shift in both tone and quality</p>
<p><b>Pace</b> (Jacobson, 1999)</p>	<p>Degree to which the pace of the program allowed for clarity and did not detract from the program.</p>	<p>Categorical: Too fast Too slow Just fine</p>

<b>Quality of the resource</b>	Degree to which the resource where program took place is awe-inspiring or particularly iconic.	3= Contextually iconic or grandiose 2= Pleasant but not iconic 1= Unimpressive/generic
<b>Unexpected negative circumstance</b> (Powell, et al., 2009)	Were there any unexpected interruptions or emergencies during the program, such as a sudden change in weather, medical emergency, technical difficulties, or hazardous conditions that detracted from the quality of the program?	1 = Yes 0 = No
<b>Unexpected positive circumstance</b>	Was there an unexpected experience that occurred during the program, such as seeing charismatic wildlife or other unique phenomena that added significantly to the quality of the experience?	1 = Yes 0 = No
<b>Behavioral theory elements</b>		
The following were only measured for programs in which a behavioral change was expressed by the interpreter as a desired program outcome. These factors follow Ajzen's Theory of Planned Behavior (1991), which suggests that a specific behavior is most influence by assessments of the costs and benefits of the expected outcomes of the behavior, social norms, and the amount of control and confidence a person has in performing that behavior.		
<b>Benefits of action</b> (Ajzen, 1991; Ham et. al., 2007; Jacobson, 1999; Knudson, et al., 2003; Moscardo, 1999; Peake et. al, 2009)	Degree to which the program emphasized the potential benefits resulting from performing a particular action(s).	4= Explicitly/purposefully emphasized 3= Mentioned a moderate amount 2= Explained a little 1= No mention NA = not applicable
<b>Costs of action</b> (Ajzen, 1991; Ham et. al., 2007; Jacobson, 1999; Knudson, et al., 2003; Moscardo, 1999; Peake et. al, 2009)	Degree to which the program emphasized the potential costs resulting from performing a particular action(s).	4= Explicitly/purposefully emphasized 3= Mentioned a moderate amount 2= Explained a little 1= No mention NA
<b>Norms of action</b> (Ajzen, 1991; Ham et. al., 2007; Jacobson, 1999; Knudson, et al., 2003; Moscardo, 1999)	Degree to which the program emphasized the social acceptability of performing a particular behavior or desired action.	4= Explicitly/purposefully emphasized 3= Mentioned a moderate amount 2= Explained a little 1= No mention NA
<b>Ease of action</b> (Ajzen, 1991; Ham et. al., 2007; Jacobson, 1999; Knudson, et al., 2003; Moscardo, 1999; Tilden, 1957)	Degree to which the program communicated the ease (or difficulty) of performing a particular behavior or desired action.	4= Explicitly/purposefully emphasized 3= Mentioned a moderate amount 2= Explained a little 1= No mention NA
<b>Demonstrates action</b> (Ajzen, 1991; Beck and Cable, 2002; Knudson, et al., 2003; Moscardo, 1999; Sharpe, 1976; Ward and Wilkinson, 2006)	Degree to which the program provided examples of, or opportunities for, performing a desired action.	4= Majority of audience engaged 3= Demonstration by ranger or small proportion of audience 2= Verbal description 1= No mention/demonstration NA

### ***Reliability and calibration***

Reliability refers to the consistency with which different researchers measured each characteristic. We built a calibration phase into the research design to ensure that each researcher's scores of each observed characteristic were consistent and reliable and therefore could be interpreted similarly. This involved three steps. First, immediately upon the completion of the field research and data entry, we carefully examined differences in the average scores of each variable between each member of the research team using a one-way ANOVA with posthoc tests. This compared each researcher's scores and identified any statistically significant differences between the mean scores for observations. Second, through detailed examination of field notes and group discussions, we determined whether any of these differences might be attributed to systematic differences in observation techniques as opposed to differences in the unique sets of programs observed by each researcher. Two types of systematic differences emerged. In the first case, one researcher was systematically higher or lower than the other three on a particular measurement scale. In these cases, scoring procedures were reviewed, consensus definitions were developed, and that one researcher re-coded the variable based on these definitions and their qualitative program notes. Variables that were re-coded in this manner included *comfort of the interpreter*, *passion*, *apparent knowledge*, *sincerity*, *provocation*, *holistic story*, and *appropriateness for the audience*. In the second case, a researcher had misinterpreted the response scale (scoring values) of the variable being coded. Again, a consensus definition was clarified and re-coding of that variable took place. These variables included *cognitive engagement*, *clear theme*, and *central message*. In one case, a variable was removed due to inconsistent interpretation of its definition in the field: *place-based messaging*.

We built in this step prior to further analysis to ensure internal consistency of the data and full confidence in the reliable measurement of each variable. This process revealed the difficulty in consistently assessing a number of subjective measures in the field, even among a well-trained team. Nevertheless, we are confident in the reliability of the data following these procedures.

### ***Data Entry and Cleaning***

Post-program surveys and program audits were coded and entered into Microsoft Access Database and Microsoft Excel to facilitate data entry. Data were then transferred to SPSS for screening and analysis. The visitor survey data were first screened for missing values and any surveys missing more than 50% of the items per factor were removed. A total of 118 respondents were removed as a result. Data were then screened for univariate and multivariate outliers on outcome variables following Tabachnick and Fidell (2007) using Mahalanobis Distance (MAH) and studentized deleted residuals (SDRESID). A total of 58 cases were removed for exceeding +/- 3 standard deviations, or the criterion Mahalanobis Distance value. This reduced our sample to 3,427 individual surveys from 376 interpretive programs.

### ***Interpretive Program Sample Development and Data Cleaning***

To answer the first research question regarding the most common practices used in live interpretation in the NPS, we used data from all 376 observed programs. However, for analyses linking program characteristics to program outcomes, some additional data cleaning was necessary. Because research question 2 uses the interpretive program as the unit of analysis, we aggregated individual data at the program level by calculating the mean score of each visitor outcome for each program. To do so, we first needed to determine how many completed surveys within a particular program would serve as a viable reflection of the quality of that program and its impacts on visitors. Prior research suggests that programs with particularly small numbers of attendees may be inherently different than programs with

larger numbers of attendees (Forist 2003; McManus 1987, 1988; Moscardo 1999). In particular, programs with fewer than five attendees may have a high likelihood of serving only a single cohesive group (e.g., a single family). Meanwhile, programs with five or more have a higher likelihood of being comprised of multiple groups. Moreover, a greater number of visitor responses enhances the reliability of the research findings. Based on this rationale, we separated programs with fewer than 5 attendees from those with 5 or more attendees, and analyzed them separately.

For groups with five or more attendees, we included in the analysis all programs with ten or more respondents to the surveys. We only included those programs with less than 10 respondents if the number of respondents represented at least half of the eligible respondents at the program (those over the age of 15). This yielded a total of 272 programs with 5 or more attendees for analysis. We employed these rules to enhance the reliability of scores for each program.

For programs with fewer than five attendees (n = 45), we only included those in which all eligible respondents to the survey (those over the age of 15) completed a survey. If a census was not achieved, the program was dropped from further analysis. This resulted in the removal of five of these smaller programs, leaving 40 in the sample for further analysis. The resulting samples are shared in Table 5.

**Table 5. Total numbers of programs and surveys used for predictive analyses, relating program and interpreter characteristics to visitor outcomes**

Park unit (total programs included in sample)	Group size	Programs	Valid surveys
Aztec Ruins National Monument (2)	< 5 attendees	2	4
	≥ 5 attendees	0	0
Badlands National Park	< 5 attendees	1	3
	≥ 5 attendees	13	115
Bryce Canyon National Park	< 5 attendees	0	0
	≥ 5 attendees	12	127
Chaco Culture National Historical Park	< 5 attendees	0	0
	≥ 5 attendees	7	70
Ford's Theater National Historic Site	< 5 attendees	0	0
	≥ 5 attendees	18	436
Fort McHenry National Monument and Historic Shrine	< 5 attendees	0	0
	≥ 5 attendees	11	113
Gettysburg National Military Park	< 5 attendees	1	2
	≥ 5 attendees	17	184
Grand Canyon National Park	< 5 attendees	1	2
	≥ 5 attendees	27	361
Great Smoky Mountains National Park	< 5 attendees	1	1
	≥ 5 attendees	11	85
Harpers Ferry National Historical Park	< 5 attendees	1	4
	≥ 5 attendees	11	75
Independence National Historical Park	< 5 attendees	2	5
	≥ 5 attendees	15	117
Jefferson National Expansion Memorial	< 5 attendees	2	6
	≥ 5 attendees	12	129

Jewel Cave National Monument	< 5 attendees	0	0
	≥ 5 attendees	18	177
Lincoln Home National Historic Site	< 5 attendees	0	0
	≥ 5 attendees	10	72
Manassas National Battlefield Park	< 5 attendees	7	18
	≥ 5 attendees	8	62
Mesa Verde National Park	< 5 attendees	0	0
	≥ 5 attendees	14	290
Mount Rushmore National Memorial	< 5 attendees	0	0
	≥ 5 attendees	9	101
National Mall	< 5 attendees	8	15
	≥ 5 attendees	8	34
Navajo National Monument	< 5 attendees	1	3
	≥ 5 attendees	2	20
Point Reyes National Seashore	< 5 attendees	5	14
	≥ 5 attendees	3	18
San Francisco Maritime National Historical Park	< 5 attendees	3	6
	≥ 5 attendees	11	58
Ulysses S. Grant National Historic Site	< 5 attendees	2	2
	≥ 5 attendees	6	34
Wind Cave National Park	< 5 attendees	0	0
	≥ 5 attendees	13	175
Yosemite National Park	< 5 attendees	3	5
	≥ 5 attendees	16	167
<b>Totals</b>	<b>&lt; 5 attendees</b>	<b>40</b>	<b>90</b>
	<b>≥ 5 attendees</b>	<b>272</b>	<b>3,090</b>
	<b>All programs</b>	<b>312</b>	<b>3,180</b>

## **Results**

### ***Data reduction and index development***

Before conducting further analyses, we conducted exploratory and confirmatory factor analyses to explore the relationships between variables and form factors made up of multiple items that represent a concept. The items that vary together as part of a factor can be combined to create scales or composite indexes that represent coherent concepts for use in subsequent analyses (DeVellis, 2003). We conducted these data reduction techniques on program observations on the full sample of observed programs. We conducted data reduction on program outcome variables using the individual respondent data after the data were first cleaned of outliers.

Exploratory factor analyses and reliability analyses on program observations revealed the presence of four latent factors: two interpreter characteristics and two program characteristics. We have named the two resulting interpreter characteristics factors *Confidence* and *Authentic emotion and charisma*. We labeled the two resulting program characteristics factors *Organization* and *Connection*. The items making up each factor are included in Table 6.

The *Confidence* factor generally reflects the notion that the interpreter appears in control of the program and is comfortable with what they are presenting. We use the term *Authentic emotion and charisma* to denote a special sort of identity that the interpreter exudes to his or her audiences. Interpreters scoring high on this factor were likeable and showed apparent and obvious passion and care for what they were interpreting. *Organization* reflects many of the best practices taught by the National Park Service's Interpretive Development Program in addition to the writings of Sam Ham (e.g., Ham, 1992). Meanwhile, *Connection* strongly reflects the core elements of Tilden's classic core principles (Tilden, 1957).

We next investigated whether the visitor outcome variables consistently varied together and potentially formed factors. Exploratory factor analyses and reliability analyses revealed the presence of two latent factors. We also ran confirmatory factor analysis, which is a form of structural equation modeling, to further develop and refine the structure of these two factors. The resulting model confirmed the two factors while also providing a more parsimonious solution. Model fit statistics were all within the acceptable range ( $S-B \chi^2=338.41$ ; CFI=.96; RMSEA=.08). We labeled the resulting factors *Visitor Experience and Appreciation* and *Behavioral Change*.

These factors form two of the three outcomes employed in this study. The third reflects attendees' satisfaction with the programs they just attended. The first factor reflects an overall assessment of the impact of the program on the individual's experience, attitudes, and knowledge. Taken as a whole, it may be the best reflection of the first two elements of the classic statement from an old NPS manual quoted by Tilden (1957), "Through interpretation, understanding; through understanding, appreciation; through appreciation, protection." The *Behavioral change* factor relates to the third part of the classic quote, actually influencing the behavior of visitors in some way. The *satisfaction* score was measured through a single survey item: "On a scale of 0 to 10, ten being the best, please rate your overall level of satisfaction with the program you just attended."

Composite indexes were created for each of the factors by equally weighting each item and taking the average of all items within the index. Table 6 shares the individual items that comprise each resulting index, as well as Cronbach's alpha scores for each. Cronbach's alpha is a measure of internal consistency of each index and can range from 0 to 1. Cronbach's alpha scores above 0.7 are considered acceptable for developing indexes (Devellis 2003). Higher Cronbach's alpha scores indicate greater internal consistency of the index.

While the factor analyses revealed that *confidence*, *authentic emotion and charisma*, *organization*, and *connection* are separate constructs, they are also moderately correlated with each other ( $r$  ranges from .357 to .623). This suggests that when an interpreter scores highly on any one of these indexes, he or she is likely to score highly on the others as well.

**Table 6. Indexes developed through exploratory and confirmatory factor analyses.**

<b>PREDICTOR VARIABLE INDEXES</b>	
<p><b>Interpreter characteristic: Confidence</b> (Cronbach's alpha = 0.70)</p> <ul style="list-style-type: none"> <li>• Comfort of the Interpreter</li> <li>• Apparent knowledge</li> <li>• Eloquence</li> </ul>	
<p><b>Interpreter characteristic: Authentic emotion and charisma</b> (Cronbach's alpha = 0.85)</p> <ul style="list-style-type: none"> <li>• Passion</li> <li>• Charisma</li> <li>• Sincerity</li> </ul>	
<p><b>Program characteristic: Organization</b> (Cronbach's alpha = 0.82)</p> <ul style="list-style-type: none"> <li>• Quality of the introduction</li> <li>• Appropriate sequence</li> <li>• Effective transitions</li> <li>• Holistic story</li> <li>• Clarity of theme</li> <li>• Link between introduction and conclusion</li> </ul>	
<p><b>Program characteristic: Connection</b> (Cronbach's alpha = 0.88)</p> <ul style="list-style-type: none"> <li>• Link to intangible meanings and universal concepts</li> <li>• Cognitive engagement</li> <li>• Relevance to audience</li> <li>• Affective messaging</li> <li>• Provocation</li> </ul>	
<b>OUTCOME INDEXES</b>	
<p><b>Program outcome: Visitor Experience and Appreciation</b> (Cronbach's alpha = 0.89)</p> <p>To what extent did the program you just attended influence any of the following for you?</p> <ul style="list-style-type: none"> <li>• Made my visit to this park more enjoyable</li> <li>• Made my visit to this park more meaningful</li> <li>• Enhanced by appreciation for this park</li> <li>• Increased my knowledge about the program's topic</li> <li>• Enhanced my appreciation for the National Park Service</li> </ul>	
<p><b>Program outcome: Behavioral change</b> (Cronbach's alpha = 0.94)</p> <p>To what extent did the program you just attended influence any of the following for you?</p> <ul style="list-style-type: none"> <li>• Changed the way I will behave while I'm in this park</li> <li>• Changed the way I will behave after I leave this park</li> </ul>	

**Descriptive statistics: Visitor characteristics**

Less than five percent of survey respondents attended the program alone. More than half (51.3%) were visiting with children. Most (59.6%) had been in the park less than one full day when they attended the program, and 37.5% had attended a ranger-led program in the same park prior to the one they were attending on the day they were surveyed. More than half of the respondents to the surveys were female (56.4%). The ages of respondents ranged from 16 to 88, with a mean of 43 and a median of 46. Eighty-nine percent of respondents described themselves as White and not of Hispanic descent. Roughly 7% described themselves as Hispanic (3.6%) or Asian (3.8%). Only 37 respondents (1.1%)

described themselves as Black and not of Hispanic descent; 15 respondents identified themselves as Native American and 26 respondents identified themselves as “other.” Twenty-seven respondents marked more than one category. Only 218 respondents (6.4%) were categorized as “local,” and 150 (4.4%) specified that they were from another country. For comparison, a 2009 survey of U.S. residents conducted by the National Park Service estimated that roughly 78% of all visitors to National Park units were White; roughly 9% were Hispanic; roughly 7% were African American; roughly 3% were Asian; and roughly 1% were Native American (Taylor et al. 2010). We were unable to locate visitor demographic information for the individual parks in this study.

**Descriptive statistics: Outcomes**

All descriptive statistics for outcomes variables were run following removal of statistical outliers. Table 7 shows attendees’ reported levels of *satisfaction* with the programs they had just attended. The mean *satisfaction* score was 8.98.<sup>3</sup>

**Table 7. Respondents’ levels of satisfaction with the program they had just attended (0 to 10 scale)**

Satisfaction (0 to 10 scale)	< 5	5	6	7	8	9	10
Percent of respondents	0.0%	1.5%	2.8%	8.2%	17.7%	22.8%	47.0%

**Table 8. Means and frequencies for items comprising the Behavioral change and Visitor experience and appreciation indexes.**

Variable (Scale)	Mean	St. Dev.	% of respondents selecting each				
			1	2	3	4	5
<b>Behavioral change (1 to 5)</b>	2.96	1.47					
• Changed the way I will behave while I’m in this park (1 to 5)	2.97	1.52	28.9	8.4	21.4	19.2	22.2
• Changed the way I will behave after I leave this park (1 to 5)	2.95	1.49	27.6	10.7	21.4	19.7	20.5
<b>Visitor experience and appreciation (1 to 5)</b>	4.43	0.65					
• Made my visit to this park more enjoyable (1 to 5)	4.55	0.70	0.2	1.3	6.7	27.0	64.7
• Made my visit to this park more meaningful (1 to 5)	4.50	0.77	0.2	2.0	8.1	26.4	63.1
• Enhanced my appreciation for this park (1 to 5)	4.38	0.79	0.5	2.1	10.0	33.8	53.5
• Increased my knowledge about the program’s topic (1 to 5)	4.46	0.77	0.4	1.9	9.1	28.6	59.9
• Enhanced my appreciation for the National Park Service (1 to 5)	4.27	0.86	1.2	2.4	13.1	34.7	48.7

<sup>3</sup> Prior to removing outliers, that mean was 8.90. Approximately 1% (n = 50) of the over three thousand program attendees that responded to the survey rated their level of satisfaction with the program they had just attended below a 5 on the 0 to 10 scale. The statistical procedures described above indicated that each of these scores below 5 was a statistical outlier. They were thus removed from the sample prior to further analyses.

The *Behavioral change* and *Visitor experience and appreciation* indexes were both measured on a scale from 1 to 5. The mean of the *Behavioral change* index was 2.96 with a standard deviation of 1.47. Forty-three percent of individuals scored above the midpoint on the scale. The mean of the *Visitor experience and appreciation* index was 4.43 with a standard deviation of 0.65. Ninety-five percent of respondents scored above the midpoint on the scale. The means and frequencies of each of the individual items associated with each index are presented in Table 8. While the *Visitor experience and appreciation* items are particularly high on the scale, items associated with *Behavioral change* were more evenly distributed.

**Question 1: What are the most common interpreter and program characteristics associated with National Park Service live interpretive programs?**

***Descriptive statistics: Program types and attendees***

We observed 376 programs. Advertised program lengths ranged from 15 minutes to 6 hours. Actual program lengths ranged from 7 minutes to 5 ½ hours. The average program length was just under 47 minutes. Two-hundred and forty-one (64%) of the programs focused primarily on cultural heritage; Eighty-eight (23%) had a primary focus on the natural environment. Thirty-nine (10%) had a dual focus and others had neither central focus (for example, general orientation talks). Programs included guided tours, talks, demonstrations, hands-on activities, and multi-media presentations. Guided tours and stationary talks made up over 80% of the programs we observed.

The number of attendees at each program ranged from 1 person to approximately 600 people. The median number of attendees was 18. Fifty-one programs (13.6%) ended with fewer attendees than they had begun with. Fifty-two programs (13.8%) were at least 20% shorter than advertised; sixty-five programs (17.3%) were at least 20% longer than advertised. Twenty (5.3%) of the programs experienced notably bad weather.

Most programs included children in their audiences. Children made up more than half of the attendees of 8% of the programs, about half the attendees on 28% of the programs, and less than half of the attendees on 46% of the programs. Only 17% of the programs had no children in their audiences.

***Descriptive statistics: Interpreter characteristics***

Three-hundred and twenty-seven of the observed interpreters were park rangers; forty-four were volunteers, and five were concessionaires. Sixty-one percent were male. Eighteen percent were under the age of 25; 23% were between the ages of 25 and 34; 26% were between the ages of 35 and 50; and 33% were over 50 years old. The interpreters averaged 9.7 years of experience in the NPS and 7.0 years in interpretation at their current park unit. Nearly one quarter of the interpreters (24.2%) had presented the program we observed at least 100 times before. More than one third (36.5%) had presented the program at least 50 times before. Nearly one-third (32.8%) had presented the program ten or fewer times. For seven interpreters this was their first time presenting the program we observed.

We asked interpreters prior to their programs to indicate the visitor outcomes that they felt were most important to achieve in that program (Table 9). Most (90%) noted more than one intended outcome. We also asked interpreters how their programs were developed (Table 10).

We asked a subset of interpreters (n = 188) about their level of excitement about the program they were about to present. The level of excitement averaged 7.81 on a ten-point scale, with responses ranging from 2 to 10 on the scale. Seven percent ranked their level of excitement below the midpoint (5) on the scale; 4% selected the midpoint; and 89% rated their level of excitement above the midpoint.

**Table 9. Intended outcomes expressed by interpreters immediately prior to their programs.**

<b>I want my audience to . . .</b>	<b>Proportion expressing each outcome</b>
Have an increased knowledge of the program topic	77.7%
Have an increased appreciation for this park	54.8%
Have an increased understanding of the park’s resources	39.6%
Want to learn more about the program topic	25.0%
Be entertained	16.2%
Have an increased appreciation of the NPS	14.4%
Have an increased concern for a specific topic	10.9%
Change their attitudes toward something	10.9%
Change a certain behavior in the future	6.6%
Develop and practice a new skill	2.9%

**Table 10. How interpretive programs were developed.**

<b>Program development</b>	<b>Proportion expressing each</b>
Program provided for ranger with full script planned out	< 1%
Program provided for ranger with some freedom to inject own style	15%
Program topic provided, little restrictions on information or style to be presented	21%
General topic suggested, but wrote own script and selected information	51%
Interpreter selected and developed entire program free of restrictions	13%

***Descriptive statistics: Interpreter delivery styles***

Tables 11 and 12 display descriptive statistics of each of the interpreter delivery styles observed in the study. Table 11 contains variables that can be considered ordinal (where variables are measured on an increasing scale). Table 12 contains binary and categorical variables, or those in which the presence or absence of the characteristics is the essential feature being measured.

***Descriptive statistics: Program characteristics***

Tables 13 and 14 display descriptive statistics for each of the program characteristics observed in the study. Table 13 displays those that can be considered ordinal variables, while Table 14 displays those that are categorical in nature.

***Qualitative descriptions of key interpreter delivery styles and program characteristics***

To provide a better sense of what particular practices looked like in action, tables that contain excerpts from the research team’s field notes, which describe examples of high and low scores of the interpreter delivery styles and program characteristics, are provided in Appendix A.

**Table 11. Descriptive statistics of interpreter delivery styles (ordinal variables).**

Variable (Scale)	Mean	St. Dev.	% of programs with each score				
			1	2	3	4	5
<b>Confidence index (1 to 4)</b>	3.26	0.49					
• Comfort of the interpreter (1 to 4)	3.47	0.61	0.3	5.3	41.7	52.7	
• Apparent knowledge (1 to 4)	3.39	0.62	0.0	7.5	46.1	46.4	
• Eloquence (1 to 4)	2.93	0.62	0.5	21.0	63.3	15.2	
<b>Authentic emotion and charisma index (1 to 5)</b>	3.52	0.82					
• Passion (1 to 5)	3.18	1.00	3.5	22.7	35.8	28.3	9.6
• Charisma (1 to 5)	3.76	0.84	0.0	7.5	27.7	45.9	18.9
• Sincerity (1 to 4)	2.89	0.75	2.7	26.6	50.3	20.5	
Responsiveness (1 to 3) <sup>a</sup>	2.81	0.42	1.2	17.0	81.9		
Humor quality (1 to 4)	2.03	0.70	21.7	55.3	21.4	1.6	
Humor quantity (1 to 5)	2.03	0.69	18.7	62.1	16.5	2.4	0.3
Personal sharing (1 to 4)	1.65	0.72	48.3	39.1	11.8	0.8	
Audibility (1 to 3)	2.85	0.37	0.5	13.6	85.9		
Formality (1 to 5)	3.20	0.83	1.1	17.3	48.1	27.9	5.6
Sarcasm (1 to 3)	1.20	0.44	81.4	17.0	1.6		
False assumptions of audience (1 to 3)	1.15	0.38	85.6	13.6	0.8		

<sup>a</sup> Responsiveness was not observable in every case. For this variable, n = 342.

**Table 12. Descriptive statistics of interpreter delivery styles (categorical variables).**

Interpreter delivery style	% of programs in which delivery style occurred
Professional appearance of the interpreter	98.4
Inequitable treatment of audience	2.9
Impatience	1.6
Primary identity: Friend	20.8
Primary identity: Authority	4.3
Primary identity: Walking encyclopedia	75.0
Character acting: partial	2.5
Character acting: complete	2.2
Interpreter bias	4.0
Questionable information	9.0

**Table 13. Descriptive statistics of program characteristics (ordinal variables).**

Variable (Scale)	Mean	St. Dev.	% of programs with each score				
			1	2	3	4	5
<b>Organization index (1 to 5)</b>	3.28	0.70					
• Quality of introduction (1 to 3)	2.08	0.45	6.1	79.5	14.4		
• Appropriate sequence (1 to 4)	2.77	0.69	2.9	29.0	55.9	12.2	
• Transitions (1 to 4)	2.68	0.75	6.6	29.8	52.9	10.6	
• Holistic story (1 to 5)	2.74	0.98	9.1	31.8	39.8	14.4	4.8
• Conclusion linked to intro (1 to 4)	2.56	0.84	10.9	34.6	42.3	12.2	
• Clear theme (1 to 4)	2.76	0.86	5.3	35.4	37.2	22.1	
<b>Connection index (1 to 5)</b>	2.73	0.73					
• Connection to intangibles/universals (1 to 5)	2.83	0.93	7.0	29.7	39.5	21.1	2.7
• Cognitive engagement (1 to 5)	2.82	0.90	4.5	34.7	36.8	21.9	2.1
• Relevance to audience (1 to 5)	2.81	0.84	3.2	35.6	39.1	21.0	1.1
• Affective messaging (1 to 5)	2.42	0.90	14.4	41.9	32.8	9.3	1.6
• Provocation (1 to 4)	2.22	0.70	13.3	54.8	29.0	2.9	
Clear message (1 to 4)	2.15	0.90	25.3	42.7	23.4	8.6	
Appropriate logistics (1 to 4)	3.11	0.92	6.9	16.8	43.8	41.5	
Appropriate for the audience (1 to 5)	3.91	0.73	0.0	3.2	21.9	56.0	18.9
Multisensory (1 to 3)	2.36	0.50	0.8	62.0	37.2		
Physical engagement (1 to 4)	1.42	0.68	67.6	25.0	5.6	1.9	
Verbal engagement (1 to 5)	2.50	0.97	16.2	34.0	34.8	13.3	1.6
Surprise (1 to 3)	1.07	0.27	93.1	6.6	0.3		
Novelty (1 to 3)	1.16	0.40	85.4	13.3	1.3		
Consistency (1 to 3)	2.87	0.38	1.6	9.8	88.6		
Resource quality (1 to 3)	2.35	0.69	12.2	41.0	46.8		
Multiple viewpoints (1 to 3) <sup>a</sup>	2.62	0.50	0.8	36.2	63.0		
<b>Behavioral theory elements<sup>b</sup></b>							
Benefits of action (1 to 4)	2.55	0.59	0.0	50.0	45.2	4.8	
Costs of action (1 to 3)	2.02	0.78	28.6	40.5	31.0		
Norms of action (1 to 3)	2.50	0.60	54.8	40.5	4.8		
Ease of action (1 to 3)	2.69	0.64	40.5	50.0	9.5		
Demonstrates action (1 to 4)	2.05	0.96	31.0	45.2	11.9	11.9	

<sup>a</sup> Multiple viewpoints were not appropriate in every case. We only observed this variable where it seemed appropriate (n = 127).

<sup>b</sup> These variables are explicitly associated with behavioral change theory. As such, they were only observed on a small subset of cases where specific behaviors were discussed by the interpreter (n = 42).

**Table 14. Descriptive statistics of program characteristics (categorical variables).**

<b>Program characteristics</b>	<b>% of programs in which characteristic occurred</b>
Fact-based messaging	24.9
Use of props	31.3
Pace too fast	5.9
Pace too slow	8.5
Pace just right	85.6
Unexpected positive circumstance	2.1
Unexpected negative circumstance	15.7

**Question 2: Which practices and approaches most consistently lead to more positive outcomes for visitors?**

***Interpreter delivery styles and program characteristics***

Table 15 displays in rank order correlations between all ordinal program characteristics and interpreter delivery styles and visitor outcomes for programs with five or more attendees. Statistical significance is displayed in two ways within the table. A single asterisk indicates that the correlation is statistically significant at  $p < 0.05$ . A double asterisk indicates that the correlation is statistically significant at  $p < 0.01$ . As such, the most powerful relationships are those with two asterisks. These are bolded and italicized for ease of interpretation. Cells with no asterisks represent no statistically significant relationships between the variables.

Behavioral theory elements were observed in 42 programs overall, including 31 with five or more attendees. Only one behavioral theory element showed a statistically significant correlation with the *Behavior change* index, “Costs of action” ( $r = .597, p < .001$ ). This suggests that programs that explicitly addressed the costs of undertaking a potential behavior were generally more successful at influencing behavior change intentions than others.

**Table 15. Pearson correlations between ordinal predictor variables and visitor outcomes for programs with five or more attendees.**

Variable	Satisfaction	Visitor Experience and Appreciation	Behavioral change
Interpreter style: Confidence index	<b>.479**</b>	<b>.277**</b>	<b>.174**</b>
Interpreter style: Authentic emotion and charisma index	<b>.423**</b>	<b>.303**</b>	<b>.182**</b>
Program characteristic: Approp. for audience	<b>.381**</b>	<b>.378**</b>	.153*
Program characteristic: Organization index	<b>.362**</b>	<b>.219**</b>	.132*
Program characteristic: Connection index	<b>.342**</b>	<b>.259**</b>	.124*
Interpreter style: Humor quality	<b>.288**</b>	<b>.233**</b>	.155*
Program characteristic: Consistency	<b>.271**</b>	<b>.281**</b>	.034
Program characteristic: Clear message	<b>.255**</b>	<b>.281**</b>	<b>.187**</b>
Interpreter style: Responsiveness	<b>.241**</b>	<b>.245**</b>	.061
Program characteristic: Verbal engagement	<b>.234**</b>	<b>.240**</b>	<b>.162**</b>
Program characteristic: Multisensory engagement	<b>.216**</b>	.115	.141*
Interpreter style: Audibility	<b>.197**</b>	.134*	.104
Interpreter style: False assumption of audience	<b>-.172**</b>	<b>-.197**</b>	-.088
Program characteristic: Appropriate logistics	<b>.170**</b>	<b>.245**</b>	<b>.165**</b>
Program characteristic: Surprise	.150*	.151*	.127*
Program characteristic: Novelty	.145*	.024	.014
Interpreter style: Humor quantity	.144*	.097	.062
Program characteristic: Physical engagement	.074	.120*	.061
Interpreter style: Formality	-.069	-.155*	-.023
Interpreter style: Sarcasm	.105	.053	-.114
Program characteristic: Quality of the resource	.077	.068	.065
Interpreter style: Personal sharing	.035	.048	.112
Program characteristic: Multiple points of view	.031	.157	.128

T-tests and ANOVAs were performed to examine the relationships of categorical variables upon visitor outcomes. These variables included *unexpected positive* and *negative consequences*, *pace*, *bias*, *impatience*, *inequitable treatment of audience*, *questionable information*, *use of props*, and interpreter identities. Table 16 summarizes only the statistically significant differences observed in the data. Pluses and minuses signify the direction of statistical significant differences. Double symbols indicate significance at the  $p \leq 0.01$  level and single symbols indicate significance at the  $p \leq 0.05$  level (see tables B-1 and B-2 in Appendix B for more detail). Programs in which the interpreter outwardly expressed impatience with the audience received lower *satisfaction* and *visitor experience and appreciation* scores than others, as did programs with an *unexpected negative occurrence*. Programs in which the interpreter employed the “friend” identity manifested higher *satisfaction* scores than others. Meanwhile, programs in which the interpreter employed the “walking encyclopedia” identity yielded lower behavioral intention scores than others. Paces that felt too fast or too slow resulted in lower *satisfaction* scores. A too slow pace was related to lower *visitor experience and appreciation* scores, and a too fast pace was associated with weaker *behavioral change* (see Table B-2). No statistically significant differences were observed for smaller programs (fewer than five attendees).

**Table 16. Significant differences in means of visitor outcome scores for selected categorical variables for programs with five or more attendees.**

Observed category	Satisfaction	Visitor experience and appreciation	Behavioral change
Impatience	-	--	
“Friend” identity	+		
“Walking encyclopedia” identity			-
Fact-based messaging	--	-	
Appropriate pace	++	++	
Unexpected negative circumstance	--	--	
The following categorical variables yielded no statistically significant differences in visitor outcomes: Inequitable treatment of the audience, questionable information, “Authority” identity, unexpected positive circumstances, use of props.			

***Program attrition and outcomes***

Program attrition (people leaving a program before it was completed) was related to both *satisfaction* and *visitor experience and appreciation* for programs with five or more attendees (see Table x), suggesting that program attrition may serve as another reasonable indicator of program quality. Thirty-six of these programs experienced attrition. The best predictors of program attrition for programs with five or more attendees included interpreters’ lack of *responsiveness* to the audience, inaudibility, *false assumptions about the audience*, the identity of the *walking encyclopedia*, inappropriate logistics, the use of props, slow pace, lack of interpreter *confidence*, a lack of *organization* of the program, and an *unexpected negative circumstance* (see Tables 17 and 18). No other interpreter or program characteristics exhibited any statistically significant relationship with program attrition at  $p < 0.05$ .

**Table 17. Independent samples t-tests comparing means of characteristics for programs that experienced attrition (people left the program early) vs. those that did not.**

<b>Characteristic</b>	<b>Program attrition?</b>	<b>Means</b>	<b>t</b>	<b>p</b>
Responsiveness of the interpreter	Yes	2.62	-2.4	.020
	No	2.83		
Audibility	Yes	2.72	-2.3	.025
	No	2.91		
False assumption of the audience	Yes	1.31	2.4	.020
	No	1.11		
Appropriate logistics	Yes	2.44	-5.0	<.001
	No	3.23		
Confidence	Yes	3.08	-2.8	.006
	No	3.32		
Organization	Yes	3.09	-2.2	.031
	No	3.36		
<b>Outcomes</b>	<b>Program attrition?</b>	<b>Means</b>	<b>T</b>	<b>p</b>
Satisfaction	Yes	8.49	-3.9	<.001
	No	9.04		
Visitor experience and appreciation	Yes	4.26	-2.6	.014
	No	4.44		
Behavioral change	Yes	2.73	-1.8	.070
	No	2.95		

**Table 18. Chi-square tests comparing programs that experience attrition vs. those that did not.**

<b>Characteristic</b>	<b>Pearson <math>\chi^2</math> statistic</b>	<b>p</b>	<b>Relation to attrition</b>
Interpreter identity: walking encyclopedia	3.6	.058	More attrition
Use of props	12.4	.001	More attrition
Slow pace	5.8	.026	More attrition
Unexpected negative occurrence	8.9	.006	More attrition

### ***Relationship between interpreter and program characteristics and outcomes in programs with less than 5 attendees***

Due at least in part to the small sample size of programs with fewer than five attendees, fewer statistically significant correlations ( $p < 0.05$ ) were observed in that sample. In rank order, they included:

Correlated with *Satisfaction*:

- Connection index:  $r = .492$ ,  $p = .001$
- Organization index:  $r = .420$ ,  $p = .007$
- Appropriate for the audience:  $.337$ ,  $p = .033$
- Humor quality:  $r = .323$ ,  $p = .045$

Correlated with *Visitor experience and appreciation*:

- Connection index:  $r = .438$ ,  $p = .005$
- Organization index:  $r = .368$ ,  $p = .020$
- Appropriate for the audience:  $.348$ ,  $p = .028$

Correlated with *Behavioral change*

- Novelty:  $r = .408$ ,  $p = .009$

Thus, a subset of the variables that predicted positive outcomes in larger programs predicted similar outcomes in smaller programs. Because only four programs within this sample experienced attrition, no additional analyses were conducted pertaining to attrition.

### ***Other key findings and contextual influences***

#### ***Group size***

Before dividing the sample based on group sizes for analyses, we first examined how these programs might be different based on our observations. Programs with fewer than five attendees were typically less formal ( $t = 2.2$ ;  $p = 0.031$ ) and more consistently appropriate for their specific audiences ( $t = 2.2$ ;  $p = 0.025$ ). Interpreters typically received lower scores on the *Confidence* index in smaller groups ( $t = 2.9$ ;  $p = 0.004$ ), as they typically appeared to be less comfortable with smaller audiences ( $t = 3.2$ ;  $p = 0.001$ ). We also less commonly observed the “friend” identity in smaller groups ( $\chi^2 = 9.5$ ;  $p = 0.002$ ). Group size, however, was not statistically related to visitor outcomes.

#### ***Program length***

No statistically significant relationships were observed between program duration and visitor outcomes for programs with five or more attendees. However, programs that were at least 20% shorter than the advertised time resulted in significantly lower *satisfaction* scores (Means: 9.01 v. 8.59;  $t = 3.9$ ;  $p < 0.001$ ). Meanwhile, programs that ran 20% or more longer than the advertised time were not statistically different from others in terms of visitor outcomes. No statistically significant trends in outcomes based on program lengths were observed in programs with fewer than five attendees.

#### ***Interpreters' background, excitement, and intentions***

For the smaller program sample (those with fewer than five attendees), no statistically significant relationships were observed between interpreter backgrounds, level of excitement, program origin, or intended outcomes and visitor outcomes.

For larger group sizes (five or more attendees), program outcomes were not related to the age, gender, or experience of interpreters, nor their degree of autonomy in program development. The interpreters' degree of excitement about the program was positively correlated with visitor *satisfaction* ( $r = .193$ ;  $p = 0.030$ ). Interpreters expressing higher degrees of excitement also exhibited higher levels of *confidence* ( $r = .222$ ,  $p = .013$ ) and *authentic emotion and charisma* ( $r = .330$ ;  $p < .001$ ). Also, volunteers tended to achieve lower degree of visitor *satisfaction* than did park rangers (means: 8.70 vs. 8.98;  $t = -2.4$ ;  $p = .019$ ).

We examined the relationships between interpreters' intended outcomes and visitor-reported outcomes by conducting independent samples t-tests. T-tests compare the means of two groups. In these cases, groups were defined by the presence of an intended outcome or not. Table 19 summarizes only the statistically significant relationships between interpreters' intended outcomes and visitor survey responses. Double symbols indicate significance at the  $p \leq 0.01$  level and single symbols indicate significance at the  $p \leq 0.05$  level (see Table B-3 for greater detail). *Visitor experience and appreciation* was the most sensitive to interpreters' intended outcomes, with five different desired outcomes related to more positive visitor responses. *Satisfaction* was related to a subset of these items. Only one intention was negatively related to visitor outcomes. Interpreters who were aiming to increase visitors' knowledge as a primary outcome of their program generally achieved lower *visitor experience and appreciation* scores. Two intended outcomes were positively related to reported behavior change by visitors: increasing the audience's level of concern and changing visitors' behaviors.

**Table 19. Significant differences in means of visitor outcome scores for selected categorical variables for programs with five or more attendees.**

Intended outcome	Satisfaction	Visitor experience and appreciation	Behavioral change
Increased knowledge		-	
Increase desire to learn	+	++	
Change attitude	+	++	
Increase appreciation for Park	++	+	
Increase understanding of resource		+	
Increase level of concern			+
Change visitor behavior			++

***The role of visitor characteristics***

We also examined the relationship between characteristics of the program attendees and their responses to the programs they attended. Older respondents exhibited higher levels of *satisfaction* ( $r = .127$ ;  $p < .001$ ) and higher scores on the *visitor experience and appreciation* index ( $r = .151$ ;  $p < .001$ ). Respondents who had been in the park longer prior to their program attendance also exhibited higher outcomes scores (Table 20). Respondents who had attended a live interpretive program in the park prior to their attendance of the program at which they were surveyed exhibited higher *satisfaction* (means: 9.07 vs. 8.8,  $t = 3.28$ ,  $p = .001$ ) and *visitor experience and appreciation* scores (means: 4.49 vs. 4.39,  $t = 4.1$ ,  $p < .001$ ) than those who did not. Female respondents exhibited higher scores on all three outcomes (*satisfaction* means: 9.04 vs. 8.92,  $t = 2.9$ ,  $p = .004$ ; *vis. exp.* means: 4.49 vs. 4.36,  $t = 5.3$ ,  $p < .001$ ; *behavioral change* means: 3.02 vs. 2.85,  $t = 3.2$ ,  $p = .001$ ). Respondents who attended the program with children exhibited more positive *behavioral change* than those who visited without

children (means: 3.07 vs. 2.85,  $t = 4.2$ ,  $p < .001$ ). White respondents who were not of Hispanic descent (89% of the sample) exhibited significantly less positive post-program *behavioral change* than non-Whites (Means: 2.92 vs. 3.44;  $t = 6.54$ ;  $p < .001$ ). Program attendees who were characterized as “local” (living within an hour’s drive from the nearest park entrance based on the map center of the zip code) exhibited significantly higher visitor experience and appreciation scores (means: 4.42 vs. 4.23,  $t = 2.4$ ,  $p = 0.016$ ).

**Table 20. One-way ANOVA comparing outcome variables for program attendees who had been in the park for different amounts of time. Items not sharing the same superscript are statistically different from one another.**

Visit length	Means		
	Satisfaction	Visitor experience and appreciation	Behavioral change
Less than one full day	8.97 <sup>A</sup>	4.42 <sup>A</sup>	2.93 <sup>A</sup>
One to two days	8.94 <sup>A</sup>	4.42 <sup>A</sup>	2.92 <sup>A</sup>
More than two days	9.10 <sup>A</sup>	4.50 <sup>B</sup>	3.20 <sup>B</sup>
Statistics	F = 2.9; p = .053	F = 4.0, p = .019	F = 7.9, p < .001

We also explored differences in predictors of outcomes in different contexts: remote vs. urban parks and different types of programs (culturally focused vs. environmentally focused). The results of these explorations are shared in Appendix C.

## **Discussion**

We set out to determine the most common practices employed within National Park Service live interpretive programs. We also aimed to determine which practices and approaches most consistently lead to more positive outcomes for visitors. Finally, we aimed to understand how the characteristics of the individuals attending the programs influence visitor responses to live interpretive programs.

The first critical lesson we uncovered is that live interpretive programs across the NPS produce very high levels of *satisfaction* in their attendees. Ninety-nine percent of program attendees rated programs as a five or better on a scale from 0 to 10, and over 85% rated programs as an 8 or better. The average *satisfaction* score was 8.9. As such, we can make two assumptions about NPS live interpretive programs relevant to our research questions. First, current practices are already achieving high degrees of visitor *satisfaction*. Second, our analyses of interpreter and program characteristics’ impacts on visitor outcomes do not separate good programs from bad programs. Rather, they help to determine which characteristics most commonly appear to make the difference between programs that achieve good vs. great visitor outcomes.

### ***Summary of the most common practices employed within NPS live interpretive programs***

Interpreters most commonly reported their intention to increase the audience’s knowledge about the program’s topic, followed closely by enhancing audience’s appreciation for the park and increasing their understanding of the park’s resources. The most common interpreter characteristics displayed during programs included *authentic emotion and charisma, confidence, audibility, and responsiveness to the audience*. The commonly observed “identity type” was that of the “walking encyclopedia.” Impatience, inequitable treatment of audience members, interpreter bias and the sharing of questionable

information were all rare within the sample. Over 98% of interpreters looked professional in their appearance. The most commonly observed program characteristics included content that was *appropriate for the audience* in attendance, *appropriate logistics* and *pace*, *multisensory engagement*, and *consistency*. *Clear organization* and *connection* to the audience were more variable. *Multiple viewpoints* were also commonly incorporated where appropriate.

### ***Summary of practices and approaches that most consistently led to more positive visitor outcomes***

Interpreters who expressed that a primary goal of their program was to increase the knowledge of the audience about their program's topic achieved lower *visitor experience and appreciation* scores than others. Those aiming to change their audience's attitudes, appreciation, understanding, and/or desire to learn achieved more positive attitudinal outcomes. Interpreters who explicitly aimed to increase their audience members' levels of concern or change their behavior were more likely to achieve more positive post-program behavioral change than others.

The best predictors of positive outcomes varied somewhat for different outcomes. *Satisfaction and visitor experience and appreciation* shared most predictors in common for programs with at least five attendees, including: *confidence, authentic emotion and charisma, appropriateness for the audience, organization, connection, humor quality, consistency, a clear message, responsiveness, verbal engagement, audibility, and appropriate logistics and pace*. *Multisensory engagement* and *fact-based messaging* (negative relationship) were additionally related to *satisfaction*. The strongest predictor of post-program behavioral change was that the program had a *clear message*.

Just over 40% of respondents expressed changes in behavioral intentions following programs. Behavioral theory suggests that interpretation (and other communication/educational experiences) should not be expected to change behavior unless a specific behavior is explicitly targeted and communication is designed to address attitudes relevant to that behavior (e.g., Ajzen, 1991; Ham et al., 2007). Less than 7% of interpreters actually listed behavior change as an explicit goal of their program. As such, a lesser percentage of positive behavioral change outcomes were expected. Because the survey items did not specify any particular behavior, we expect that the attitudes addressed in programs may often be relevant to behaviors, whether an explicit goal of programs or not. Such behaviors might be as simple as taking the initiative to learn more about a topic or visiting another park site in the future or they could be more directly related to stewardship of the resource and support of the NPS mission. In either case, programs in which the interpreter explicitly targeted behavior change as an intended outcome were more successful at doing so. For a broader discussion of behavior change and interpretation see Ham et al. (2007) and Ham (2009).

A smaller subset of interpreter and program characteristics were correlated with outcomes for smaller programs (those with fewer than five attendees). *Connection, organization, and appropriateness for the audience* were each correlated with *satisfaction* and *visitor experience and appreciation*. *Humor quality* was additionally correlated with *satisfaction*. Only *novelty* was correlated with post-program behavioral change for these smaller programs.

### ***Different outcomes for different visitor types***

Attendees who were older, local, female, had been in the park longer, had attended a live interpretive program in the park before, or were attending with children generally expressed more positive outcomes. The findings suggest that certain respondents may be more receptive to interpretive

messaging than others and that cumulative experiences within a park may enhance positive impacts. Moreover, programs are often particularly well-suited for groups with children.

The behavioral intentions of non-White program attendees were more positively affected than those of White attendees. However, the sample numbers of non-White program attendees were strikingly low. Our field observations suggest that non-White attendees were no less likely to respond to our surveys than White attendees. Rather, very few non-White people appear to attend these programs. In the absence of demographic data for all visitors to any of the park units in the study, we cannot determine whether program attendance patterns are any different from park attendance patterns.

## **Conclusions and recommendations**

One of the primary objectives of the overall research was to develop a list of best practices to inform interpretive training, assessment, and monitoring in the National Park Service. This study not only informs the identification of those best practices, but also provides insights on how monitoring might most effectively take place in the NPS in the future.

### ***Best practices across all programs***

The following is a list of practices related to positive visitor outcomes that cut across contexts. While *humor quality* also was positively related to outcomes, we don't feel it constitutes a best practice, as not all programs *should* necessarily be funny, nor is humor typically a trainable item.

1. Confidence
  - Comfort, eloquence, apparent knowledge
2. Authentic emotion and charisma
  - Passion, sincerity, charisma
3. Appropriateness for audience
4. Organization
  - Quality of introduction, appropriate sequence, effective transitions, holistic story, clear theme, link between introduction and conclusion
5. Connection
  - Links to intangibles and universal concepts, cognitive engagement, relevance to audience, affective messaging, provocation
6. Consistency
7. Clear message
8. Responsiveness
9. Audibility
10. Appropriate logistics
11. Verbal engagement
12. Multisensory engagement
13. Appropriate pace
14. Avoid focusing on knowledge gain as the program's central goal and communicating solely factual information
15. Avoid making uncertain assumptions about the audience

Although each of these practices was statistically correlated with better outcomes, variability within the sample suggests that the entire suite of best practices is not a necessary precursor to a high quality program. Rather, each of these practices in various combinations was found to enhance outcomes across a majority of programs in which they were practiced. A wide range of diverse approaches led to positive visitor outcomes. As such, we recommend maintaining the freedom for interpreters to be creative and innovative in their presentations.

### ***Monitoring***

A major goal of this research was to identify best practices that might allow for the efficient monitoring of program quality. The results of our study suggest that visitor outcomes were consistently high (above the midpoint) on almost all programs in the study. This suggests that most programs are having a positive impact on visitors' experience regardless of the intensity and quality of these programs and the individual best practices employed. The best practices listed above rather help to distinguish between good and great programs. As such, simply tracking the numbers of programs and attendees provides a reasonable method for gauging the overall impacts of parks' interpretive programs on visitors. Roughly 95-99% of attendees have a positive experience at these programs, regardless of their quality as measured within this study. As such, we can assume that, to some extent, any program enhances the visitor experience in some way.

Our research also suggests a few additional indicators for program quality that may be efficiently monitored. Of the programs we attempted to attend, only 77% actually took place. In only a few cases, bad weather or other extenuating circumstances prevented a program from occurring. In others, either the interpreter or an audience did not show up. Given that most programs produce positive impacts on visitors, this may in fact be the most powerful measure of effectiveness of a park unit's interpretive programs. Low scores could reveal problems with developing, scheduling, and/or marketing interpretive programs that could directly be addressed if monitored.

As the NPS aspires to have not just good but excellent programs, monitoring of specific best practices as indicators of quality might also be appropriate. This research suggests that certain best practices are consistently linked to desired outcomes across a range of contexts. However, this research also revealed that even a well-trained team needs regular reliability checks and calibration to ensure consistent measurement. As such, we believe it to be highly unlikely that individuals at the park level who are not regularly working together would maintain consistent measurement for valid comparisons across park units. Ideally, an independent team of experts would assess interpretive programs at each park, providing consistent and unbiased measures of key characteristics.

For monitoring interpretation service-wide, our recommendation is to monitor numbers of programs and attendees, as well as the proportion of scheduled programs that actually take place. Because nearly all programs produce positive results and these data have the highest potential to be measured consistently, they serve as the most reliable measures of interpretive program health across parks. Additional measures that address audiences reached and messages communicated may also be warranted. We recommend that quality measures be used to inform training within the NPS. Unless resources are made available to develop a consistent auditing team within the agency, we do not recommend using within-park monitoring results of these best practices to compare unit to unit.

### ***Adaptive management at the park level***

Tracking of the use of best practices could potentially be incorporated into current staff evaluation and park-level monitoring systems. As supervisory rangers conduct program observations, they could address the characteristics suggested above to conduct quality assessments. These assessments could then be used as feedback for interpretive staff. However, we do not recommend using such measurements for comparisons between parks (see discussion above).

### ***Training***

We feel that the biggest contribution of this study is the validation (and in some cases identification) of practices and delivery styles linked to enhanced visitor outcomes. We urge the NPS to focus upon the list of these practices shared above to enhance current interpretive training within the NPS.

### ***Specific management implications and recommendations***

1. NPS interpretive programs tend to produce high levels of satisfaction and experience enhancement for attendees regardless of their specific characteristics.
2. This study identified a set of practices that appear to commonly make the difference between good programs and excellent programs (with regard to visitor outcomes). We recommend these practices be incorporated into current interpretive training in the NPS where appropriate.
3. The results of our study provide a holistic perspective on practices related to positive outcomes for program attendees. However, each of these practices employed in isolation does not guarantee a high quality program. Our study instead supports the notion that interpretation is a complex phenomenon that requires competence in a range of techniques and approaches that should be responsive to different audiences and contexts. As such, we recommend maintaining the freedom for interpreters to be creative and innovative in their presentations.
4. It required considerable and iterative training, feedback, and adjustment for our team to produce consistent and reliable monitoring results. Our experiences suggest that any efforts to monitor program quality across parks will require a highly trained team with consistent membership that is external to the parks.
5. Because nearly all programs produce positive results and these data have the highest potential to be measured consistently, we recommend monitoring numbers of programs and attendees, as well as the proportion of scheduled programs that actually take place. These appear to be the most reliable measures of interpretive program health across parks. Additional measures that address audiences reached and messages communicated may also be warranted. Unless resources are made available to develop an external auditing team within the agency, we do not recommend using monitoring of these best practices to compare unit to unit.
6. We recommend that the best practices uncovered in this research be used to inform interpretive training within the NPS and by interpretive staff within parks to support reflective learning and adaptive management (iterative improvement) at the park level.

### ***Limitations***

1. The treatment (an interpretive program in a National Park setting) is a complex phenomenon that is influenced by an interaction between the resource and its qualities, the social environment, including the make-up of social groups, the characteristics of the interpreter and the individual attendees, and the topic and characteristics of the program. This research examines only the relationships between visitor outcomes and selected interpreter and program characteristics. As such, other potential influences are not accounted for.
2. The results of this study may be influenced by a “ceiling effect,” which describes the phenomenon when individuals (in this case, NPS visitors) come into an experience with already high scores on the outcomes considered (in this case the specific attitudes and intentions measured in the study). As such, some respondents would report little to no change for an outcome measure because their attitudes or intentions may already be at the high end of the spectrum for the outcome in question. In these cases, the survey items may not be sensitive enough to detect the influence of a program.

3. To be able to compare responses across a wide variety of programs, visitor survey items were designed to be rather general in their content. As such, some may not reflect the content or intended outcomes of particular programs.
4. Self-reported behaviors and behavioral intentions may not reflect actual behaviors.
5. Social desirability bias, or the desire of respondents to provide answers that are socially acceptable, may influence the results of this study. This has a tendency to inflate positive findings.
6. The *Hawthorne effect* describes the phenomenon by which the presence of the researcher influences the nature of the phenomenon under observation or the results of the study (Babbie, 2007). Participating interpreters were aware that their programs were being observed by a researcher which may have altered how they conducted their programs.

## **Literature Cited**

- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Decision-Making Processes*, 50(2), 179-211.
- Babbie, E. (2007). *The Practice of Social Research*. Belmont, CA, Wadsworth.
- Beck, L., & Cable, T. T. (2002). *Interpretation for the 21st century: Fifteen guiding principles for interpreting nature and culture* (2nd ed.). Champaign: Sagamore.
- Brochu, L., & Merriman, T. (2002). *Personal interpretation: Connecting your audience to heritage resources*: InterpPress.
- Eppley Institute for Parks and Public Lands (2012). "About interpretation." <http://idp.eppley.org/about-interpretation>. Last accessed July 16, 2012.
- Forist, B. (2003). *Visitor Use and Evaluation of Interpretive Media. A Report on Visitors to the National Park System*. National Park Service Visitor Services Project. [http://nature.nps.gov/socialscience/docs/Visitor\\_Use\\_and\\_Evaluation.pdf](http://nature.nps.gov/socialscience/docs/Visitor_Use_and_Evaluation.pdf). Accessed May 24, 2012.
- Ham, S. H. (1992). *Environmental interpretation: A practical guide for people with big ideas and small budgets*. Golden: Fulcrum.
- Ham, S. (2009). From interpretation to protection: is there a theoretical basis? *Journal of Interpretation Research* 14(2): 49-57.
- Ham, S. H., Brown, T. J., Curtis, J., Weiler, B., Hughes, M., & Poll, M. (2007). *Promoting persuasion in protected areas: A guide for managers. Developing strategic communication to influence visitor behavior*. Southport, Queensland, Australia: Sustainable Tourism Cooperative Research Centre.
- Ham, S. H. & Weiler, B.M. (2002). Tour guide training: a model for sustainable capacity building in developing countries. *Journal of Sustainable Tourism* 10(1): 52-69.
- Interpretive Development Program (2008). National Park Service: Professional Standards for Learning and Performance. U.S. Department of the Interior.
- Jacobson, S. K. (1999). *Communication skills for conservation professionals*. Washington, D.C.: Island Press.
- Knapp, D., & Benton, G. M. (2004). Elements to successful interpretation: A multiple case study of five national parks. *Journal of Interpretation Research*, 9(2), 9-25.
- Knapp, D. & Yang, L. (2002). A Phenomenological Analysis of Long-Term Recollections of an Interpretive Program. *Journal of Interpretation Research*, 7(2), 7-17.
- Knudson, D. M., Cable, T. T., & Beck, L. (2003). *Interpretation of cultural and natural resources* (2nd ed.). State College: Venture Publishing.
- Larsen, D. L. (2003). *Meaningful Interpretation: How to Connect Hearts and Minds to Places, Objects, and Other Resources*. Eastern National.
- Lewis, W. J. (2005). *Interpreting for park visitors* (9th ed.). Fort Washington: Eastern National.
- Madin, E. M. P., & Fenton, D. M. (2004). Environmental interpretation in the Great Barrier Reef Marine Park: An assessment of programme effectiveness. *Journal of Sustainable Tourism*, 12(2), 121-137.
- McManus, P.M. (1987). It's the company you keep . . . The social determination of learning-related behaviour in a science museum. *The International Journal of Museum management and Curatorship* 6: 263-270.

- McManus, P.M. (1988). Good companions. More on the social determination of learning-related behaviour in a science museum. *The International Journal of Museum management and Curatorship* 7: 37-44.
- Moscardo, G. (1999). *Making visitors mindful: Principles for creating quality sustainable visitor experiences through effective communication*. Champaign: Sagamore.
- National Association for Interpretation (NAI) 2012. Mission, Vision and Core Values. [http://www.interpnet.com/about\\_nai/mission.shtml](http://www.interpnet.com/about_nai/mission.shtml). Accessed May 24, 2012.
- Powell, R.B., Skibins, J.C. & Stern, M.J. (2010) *Linking interpretation best practices with outcomes: A review of literature*. Clemson University and U.S. National Park Service, National Education Council.
- Powell, R. B., Kellert, S. R., & Ham, S. H. (2009). Interactional theory and the sustainable nature-based tourism experience. *Society & Natural Resources*, 22(8), 761-776.
- Regnier, K., Gross, M., & Zimmerman, R. (1992). *The Interpreter's Guidebook: Techniques for Programs and Presentations*. Interpreter's Handbook Series. Stevens Point: UW-SP Foundation Press.
- Skibins, J.C., Powell, R.B., and M.J. Stern (2012). Exploring empirical support for interpretation's best practices. *Journal of Interpretation Research*.
- Sharpe, G. W. (1976). *Interpreting the environment*. New York: John Wiley & Sons.
- Stern, M.J. and R.B. Powell 2011. *The Views of National Park Service Superintendents and Interpretation and Education Supervisors on Interpretation and Education in the National Park Service*. Submitted to the National Education Council of the U.S. National Park Service. October, 2011.
- Taylor, P.A., Grandjean, B.D., Gramann, J.H. (2011). *National Park Service Comprehensive Survey of the American Public 2008-2009: Racial and Ethnic Diversity of National Park System Visitors and Non-Visitors*. Natural Resource Report NPS/NRSS/SSD/NRR-2011/432. Wyoming Survey & Analysis Center, University of Wyoming, Laramie, Wyoming. Online: [http://nature.nps.gov/socialscience/docs/CompSurvey2008\\_2009RaceEthnicity.pdf](http://nature.nps.gov/socialscience/docs/CompSurvey2008_2009RaceEthnicity.pdf)
- Tilden, F. (1957). *Interpreting our heritage* (3rd ed.). Chapel Hill: The University of North Carolina Press.
- Veverka, J. A. (1998). *Interpretive master planning: the essential planning guide for interpretive centers, parks, self guided trails, historic sites, zoos, exhibits and programs* (2nd ed.). Tustin: Acorn Naturalists.
- Wallace, G. N., & Gaudry, C. J. (2002). An evaluation of the "authority of the resource" interpretive technique by rangers in eight wilderness/backcountry areas. *Journal of Interpretation Research*, 7, 43-68.
- Wearing, S., & Wearing, B. (2001). Conceptualizing the selves of tourism. *Leisure Studies*, 7, 11-23.
- Ward, C.W., & Wilkinson, A. E. (2006). *Conducting meaningful interpretation: A field guide for success*. Golden: Fulcrum.

**APPENDIX A. QUALITATIVE EXAMPLES OF KEY VARIABLES FROM THE FIELD**

**Table A – 1. Qualitative field notes describing interpreter characteristics observed during programs.**

<b>Variables</b>	<b>Examples</b>
<b>Comfort of the Interpreter</b>	<p>HIGH: The ranger used a very conversational tone when interacting with the audience. At each stop he would sit down on a fence post or lean against a sign while continuing his story. He asked visitors to stop him with questions and to suggest answers to various questions he posed. He asked the audience to stop him if they needed anything and never seemed to get distracted from his intended performance.</p> <p>- Despite numerous interruptions and difficult questions from visitors, the ranger handled each situation with ease and was not swayed from presenting their program. He encouraged visitors to challenge him and take part in the “experience,” rather than just sitting back and listening to him talk.</p> <p>LOW: Ranger was clearly unnerved by a large crowd consisting of a mix of adults and very distracted children who were bored by the historical topic of the talk. He mentioned that Civil War history was not his area of expertise and struggled to remember certain numbers and facts. He was unable to answer most visitors’ questions and did not maintain the large group very well when moving from location to location. He tried several times to stop visitors from leaving the program and looked clearly saddened each time more people left.</p> <p>- The ranger seemed very nervous and was visibly shaking, had to pause several times to collect thoughts and recall what they intended to say next. They apologized frequently for forgetting what they had scripted and relied on “um, yeah, and like” to fill in the gaps.</p>
<b>Apparent Knowledge</b>	<p>HIGH: This volunteer knew more about plants than just about anyone I've ever met. Not only did she know facts and scientific details about every plant, but also stories about their connection to humans and how people have used them in the past. She answered every single question posed by visitors, including Latin names, habitat ranges, and various vascular functions. She never seemed like she had to think before answering and was confident in every response she gave.</p> <p>LOW: The ranger attempted to tell us the name of the man who designed the memorial, the date it was commissioned, and who funded its construction, but could not remember any of these things. He referred to his notes continually throughout the program and sometimes spent an extended period of time looking through them to try to recall the fact he had intended to give. When visitors asked questions, he would again refer to his notes and generally still could not provide an answer.</p> <p>- Ranger mentioned halfway through the program that it was her first time giving it, which was evidenced by her difficulty recalling facts/figures, her regular use of notes, and long walks between stops without talking to visitors at all (while she reviewed her notes).</p>
<b>Eloquence</b>	<p>HIGH: Each story told by the interpreter was clearly illustrated through a strong vocabulary and a purposeful use of words. Pauses were only used when necessary for effect and the ranger never seemed unsure of what to</p>

Appendix A. Qualitative field notes.

	<p>say next. Their manner of speaking was concise and to the point, but did not feel rehearsed or overly scripted.</p>
	<p>LOW: Ranger said “like” often and used “um” as filler every time he paused or tried to think of an answer. He constantly used the phrase “y’know,” followed by a long pause at times when the audience was completely unsure of what he was talking about. He mumbled at times when he didn’t seem confident in what he was saying. This left the visitors feeling very confused.</p>
<p><b>Passion</b></p>	<p>HIGH: The ranger explicitly told us that he was excited to share information with us about the natural resources found within the park. He said things like “let me tell you why I love this plant so much” and “I bet you can see why this is such a cool place.” He had the audience look at things and feel them, tell the group what they liked best about it, and share their own reasons why the park was so special to them.</p> <p>- This ranger told us why the park makes him feel inspired, what he loves most about it, and makes him come alive. He had us reflect on our own feelings about the place by sharing stories. He jumped from rock to rock with an obvious excitement in his step and couldn’t wait to share his next story. When the topic called for a more somber and reflective tone he slowed down, removed his hat, and reminded us why we should care about this place.</p> <p>LOW: This ranger had a very calm, quiet demeanor that she used to convey facts about the battles that unfolded in the park. Her tone of voice and slow pace made it sound like she didn’t really care what she was saying to the audience. At one point she said, “the Civil War isn’t really my area of expertise, but it’s worth knowing something about.” Her connection to the topic and concern for preserving the place were obviously low. She would point out things along the way and say “I think this is where ____ happened” or “some people find this interesting.”</p>
<p><b>Charisma</b></p>	<p>HIGH: Ranger was kind and smiling. Imagine sweet grandmother figure telling stories about history and her childhood. Audience leaned in to hear what she had to say and observe what she was doing. Both the interpreter and audience had smiles on their faces throughout the program.</p> <p>- The ranger had a deep laugh that put smiles on the faces of visitors. He used friendly, casual banter throughout the program to keep visitors engaged and to inquire about their specific interests and hobbies. Visitors seemed to enjoy walking around with the ranger and hearing the information he had to share.</p> <p>LOW: Ranger had a very abrupt manner of speaking to visitors and sounded annoyed to have them on the program. When some questions were asked, he ignored them entirely and hurried along with the program. He made no effort to engage the audience or carry on a conversation, rather he seemed focused on presenting what he had prepared and getting away from visitors as soon as he was finished.</p>
<p><b>Sincerity</b></p>	<p>HIGH: While leading a tour of a war memorial, this ranger maintained a very solemn and respectful demeanor. He told us about the hard work, sacrifice, and heartache of people at home and abroad that made the war effort possible. Upon entering the memorial, he removed his hat and stood silently for a moment to take it all in. As he talked about each feature of the memorial he would touch it gently and slowly shake his head. His emotional</p>

Appendix A. Qualitative field notes.

	<p>connection to both the resource and the topic were clear.</p> <p>LOW: This ranger spoke in a very monotone, droning manner. At each stop, she listed several facts and then moved on to the next stop. She didn't wait for visitors to observe or enjoy the various resources and seemed to have no interest in looking at it herself. She seemed bored with the experience and made us feel like there was nothing really special about the place. Her cold delivery of facts and numbers about the battle that took place there made her seem almost callous to the topic.</p>
<b>Inequity</b>	<p>PRESENCE: At one point, the ranger entered a small room in the ruins with a big step down, and one elderly lady didn't want to try entering. She had to wait outside the room while he continued the program for another twenty minutes--only five of which actually had to do with the room we were in. The lady left with her husband after we exited the room.</p> <p>ABSENCE: The audience consisted of a wide mix of children, families, and older adults. Each time the ranger told a story or told us how something works, he would say "those of you who have kids know that" or "I bet you guys like to play with toys too." No matter what the topic was, he made sure to connect it directly to someone in the audience. When props were passed around, he checked to make sure that everyone got the chance to see it. He encouraged the group to stand so that shorter members were in the front and everyone could see what we were looking at.</p>
<b>Humor Quality</b>	<p>HIGH: The ranger poked fun at the notorious love life of a Civil War general to bring some welcome humor into an otherwise serious talk. He told us about pranks that soldiers would play on one another and had us laughing. This helped the program not only to avoid being far too sad/somber, but also connected us with the fact that these were regular people just like us.</p> <p>LOW: The ranger tried to use corny jokes and silly metaphors throughout the program to get laughs out of the audience. For the most part, the audience did not seem to find these funny. He relied so heavily on these jokes that the rest of his program was largely devoid of worthwhile information. The audience seemed tired and uninterested by the end of the program, but he kept cracking bad jokes anyway.</p>
<b>Personal Sharing</b>	<p>HIGH: The ranger shared songs he had personally written about Tuolumne Meadows, Yosemite, and the bird that was the topic of the talk. He told us about how Yosemite had a special place in his heart and how it had played a role in shaping his life.</p> <p>- This interpreter grew up not far from where the program took place and knew a great deal about the people who had lived there. She knew some of the people personally and shared several stories from her own childhood. She told us how she had gotten involved volunteering in the park and why it was worth it to her to donate her time.</p> <p>LOW: The ranger began the program without introducing himself to the audience. During the program he did not share any personal information, opinions, or connections to his stories. He shared a number of facts about the buildings we passed and told us stories about the people who lived there. We knew nothing about why he worked in the park or why he bothered telling us about these places.</p>
<b>Formality</b>	<p>HIGH: The ranger referred often to his position as a park ranger (his authority) and told visitors that it was his responsibility to make sure that visitors were safe and accounted for. He reminded the group that he wanted us to have a good time, but that our safety and understanding of</p>

Appendix A. Qualitative field notes.

	<p>the rules came first. He used his authority to guide visitors through a very crowded urban park, without getting separated or losing the flow of the program.</p> <p>LOW: Because only two other visitors showed up to the program, the ranger spoke directly to them in a very conversational tone. He asked them what they were interested in, what they would like to do on the program, and told them to stop him whenever they wanted to ask a question. In this way, the talk was almost completely catered to the interests of the visitors and they felt very comfortable even in such a small group.</p> <p>- The ranger remained seated through much of the program and did not seem to have any clear plan for the program. He allowed visitors to ask questions, but after answering each question did not provide any more content. He simply sat and waited for visitors to ask another question. Advertisement for this program indicated it would be a talk given by a ranger, but he treated it much more like an informal question and answer session, which caused many members of the audience to leave early.</p>
<p><b>Sarcasm</b></p>	<p>HIGH: Ranger kept comparing her general lack of skill/knowledge on the subject to the native people who were so talented in these areas, to illustrate just how expert these early people were at reading and understanding the land. She poked fun at herself for being “outdoors illiterate” and “as useful on a deserted island as a rock,” but kept the audience laughing and entertained.</p> <p>- The interpreter used a very negative type of sarcasm throughout the program that gave it a very pessimistic undertone. He kept referring to the “average visitor” and their general lack of knowledge about park rules, funding, or the reasons behind preservation. He would say things such as “I’m sure <i>you</i> guys would never do that” and “this group is smarter than that though, <i>right?</i>” It seemed his intention was to illustrate why it’s important to understand these things, but it came across as an implication of how dumb the audience members on his program were.</p>

Appendix A. Qualitative field notes.

<p><b>Bias with Negative Impact</b></p>	<p>PRESENCE: During a talk about global warming, the ranger spent much of the time talking about what the government was/wasn't doing (including a few potentially politically inflammatory statements), at which point several visitors walked away from the program. He remarked that it was the fault of big industry that the environment was in such bad shape and that it would take strong government regulation to reverse the trends. Most of these opinions were presented as if they were facts.</p> <p>- This ranger tended to deliver his own opinions as if they were facts. He also tended to assume that audience members agreed with his point of view and seemed to offend a couple of visitors from the Deep South. He made comments about the North being the "good guys" and the South being the "enemy." He had very strong opinions about who was/wasn't a hero here, how Hollywood has dramatized the events, and about who deserves to win a Medal of Honor (now and in the past). This also seemed to bother a couple of people. The Gettysburg Bike Week was in full swing, which made it hard to hear and he was clearly annoyed/spoke badly about them, even though there were bikers on the program. The contrast between his generally great personality/performance and his patience/tact with handling other aspects of the program were surprising. He also closed with a very conservative, patriotic message and a "god bless you, God bless America" that seemed questionably religious.</p>
<p><b>False Attribution</b></p>	<p>PRESENCE: This ranger told the audience that with the exception of a few families, "people were glad that the park had been created here." Without any substantiation, she claimed that residents who had to be removed when the park was created were fairly compensated and had a better standard of living after moving. There was no evidence provided to support this claim.</p> <p>ABSENCE: This ranger did an exceptional job of portraying all sides of the history here and speaking of soldiers and officers on both sides of the war as people just like you and me. Both Union and Confederate soldiers were included in the stories, their sides were portrayed without opinion or bias, and neither was given the image of being the "good guy" or the "bad guy" in the battles. He did not speak disparagingly about either side and did not place blame for the Civil War on any one group.</p>
<p><b>False Assumption of Audience</b></p>	<p>PRESENCE: During the program, the ranger regularly referred to names and dates very specific to events during the Civil War. These were used without any further explanation and assumed that the audience already had a fairly thorough knowledge of the Civil War. There was a small group of war "buffs" who seemed to follow and enjoy the program, but most of the rest of the audience seemed somewhat lost and disconnected without this extra knowledge.</p>

**Table A – 2. Qualitative field notes describing interpreter characteristics observed during programs.**

<p><b>Intro Quality</b></p>	<p>HIGH: Ranger began the program by saying “It is the morning of the first battle of Manassas. It’s hot and muggy. You’ve just finished breakfast, and you’re preparing for a long march over these fields you see before you. But before the day is done, half of your company will be brought down by confederate cannon and musket fire...” This captured our attention, set the tone for the program, and led directly into the theme of the program.</p> <p>LOW: The ranger arrived just in time to start the program and did not interact with the audience at all or provide any information about the program before it started. The first thing he said to the audience was “OK, let’s get started,” at which point he walked off to our first stop. When we arrived at the first stop, while much of the group was still walking, he started talking about trees and listing facts about them. There was really no introduction to the talk, nothing to capture our attention, and nothing to let us know that we were even on the right program.</p> <p>LOW: The ranger did not give any orientation as to what we would be doing on the program. When asked by an older couple what they could expect on our walk, he simply said “You’ll just have to come along and find out.” While he said this with a smile on his face, it clearly left the guests feeling unnerved and they left the group before the first stop, at which point he said “Oh, I guess they decided not to join us.”</p>
<p><b>Appropriate Logistics</b></p>	<p>HIGH: The ranger arrived before the program was scheduled to begin and announced several times what the program was and when it would be starting. This gave everyone the chance to get ready and know that they were in the right place. Once the program began, the ranger let the audience know how long we would be gone, what we would be doing, and what supplies they should have. He reminded everyone to use the bathroom before we went out on the trail and to wear sunscreen. Once on the trail, he made sure to keep the group together and maintain a reasonable pace. We stopped at spots along the trail that were out of the way of other hikers, quiet, and cool. Once the program ended, he walked with the group back to where we had started.</p> <p>LOW: The ranger showed up to this program three minutes after its designated start time. He told the group that it was his first time ever giving it and that he wasn’t sure exactly what we were supposed to be doing. The program was scheduled for an hour, but only lasted 30 minutes. The tour only had two stops, one at the parking lot and one about 100 yards away, even though it was advertised as a walking tour.</p> <p>LOW: During the walk, we stopped at a historical structure and the ranger allowed the group to explore inside the building and around the grounds for an extended period of time. This broke up the flow of the program and left 15-20 people behind as we moved on to the next spot. The ranger made very little effort to round up the group and did not announce when we would be leaving.</p> <p>LOW: Ranger kept the audience in the very hot sun when he could easily have led to cooler, shadier rooms in the ruins.</p> <p>LOW: During one of the hottest days of the entire summer, the ranger led us on a guided tour through an outdoor memorial. Even though there were opportunities for shade and</p>

Appendix A. Qualitative field notes.

	<p>places to sit down, he left us standing in direct sunlight for extended periods of time. We were all clearly drenched in sweat and moving slowly, but he continued with the program without any offer of shade, water, or assistance.</p>
<p><b>Appropriate for the Audience</b></p>	<p>HIGH: The audience at this campfire program consisted mainly of families, along with a few older adults. The program was very family friendly, with songs and activities that everyone could enjoy. There was also a great deal of content that was relatable to a young audience, but that also taught the adults things they didn't know. For parts of the program that were rather silly, adults were given roles helping to guide the kids and be in charge.</p> <p>LOW: There was only one woman with two very young kids on the tour. The interpreter did not adapt the program at all to the kids and instead seemed impatient when one was running around. She dealt with the matter by picking up the child and holding her.</p> <p>LOW: Some gory descriptions of Civil War soldiers, their injuries, and medical treatments of the time period may have been too graphic for some of the younger children in the audience.</p> <p>LOW: Although the audiences consisted of a dozen adults and only one child, the interpreter spent the entire program speaking only to the child. He used very basic language and got down on one knee to tell her certain things. This was certainly a great experience for the child, but left the rest of the group wanting more. The program was advertised as a history of FDR's life and his role in preserving the United States during war and economic depression, but everything was limited to a very basic level.</p>
<p><b>Appropriate Sequence</b></p>	<p>HIGH: This program was about the life cycle of a giant sequoia tree. The program itself followed a storyline that described the life of a tree and everything it saw during its lifespan. Each stop was related to the next stage of life and provided a clear example of that stage. We moved from an area full of cones and seeds, to a stop with several tiny saplings, to young trees, and on up to full size giants. We followed the growth of a sequoia from birth to death and understood everything it must overcome in the process.</p> <p>HIGH: Interpreter used his sequencing to enhance his messaging by moving through animals in different eco-regions of the park, but also connecting them through the food chain (food chain is also part of his transitions).</p> <p>LOW: This talk provided a random assortment of facts and stories about both the War of 1812 and the Civil War. Each stop was disconnected from the next and jumped back and forth between the two wars. There was no logical sequence to the stops and seemed to be representative of whatever was on the ranger's mind at the time. At a single stop we talked about iron clad battleships during the Civil War and a tavern that was located on the grounds during the War of 1812, with no connection drawn between them or any of the other stops.</p>
<p><b>Transitions</b></p>	<p>HIGH: As we prepared to leave each stop, the interpreter said "I want you to be on the lookout for _____ as we head for our next stop and think about how it relates to _____." This kept the visitors curious, engaged, and thinking about the theme of the talk even while the ranger wasn't talking. These transitions provided a logical flow from the topic of one stop to the next.</p>

Appendix A. Qualitative field notes.

	<p>LOW: At each stop, the ranger would talk for a bit and then just stop. We would walk to the next stop in silence and then he would pick up right where he left off. It felt very much as if he were stopping halfway through a paragraph, waiting a bit, and then continuing without any explanation of why we had moved. It likely would have been more effective to just stay in one place and deliver a talk, as these long pauses left the audience bored and distracted from the program itself.</p> <p>LOW: Interpreter left long awkward gaps with no talking and then would randomly start up again.</p>
<p><b>Physical Engagement</b></p>	<p>HIGH: Mt. Rushmore was a park that was difficult to create physical engagement with, but one ranger told people to feel the sidewalk because that's how smooth the carved faces of the presidents are.</p> <p>HIGH: Had visitors actually march in formation like soldiers would have, following the same path used in the Civil War.</p> <p>HIGH: Ranger brought out model ships, flags, etc. and let kids move them around on a large map on the ground to show where different troops were during different battles.</p> <p>HIGH: Ranger let one child try on replica Civil War gear and mimic the motions of loading and shooting a rifle. Also passed around replica soldier uniform and bullets for visitors to feel and get a sense for their weight.</p> <p>HIGH: Had visitors do traditional exercise involving circles and hand motions before entering sacred round house, as Miwok tribe members would have done before entering.</p>
<p><b>Verbal Engagement</b></p>	<p>HIGH: After sharing and explaining different sets of data on the giant video sphere, the rest of the program was treated like a discussion session with the audience members talking about what may be causing trends in climate change and how the trends may be reversed.</p> <p>HIGH: Visitors sang along with campfire songs, answered questions, and were allowed to tell stories of their experiences in the park.</p> <p>HIGH: Had visitors participate in an exercise similar to what schoolchildren would have done in the schoolhouse where the program took place. We answered questions and repeated lessons back to the "teacher."</p> <p>LOW: Ranger occasionally struggled for words and asked many rhetorical questions that didn't encourage visitor involvement. Eventually I stopped thinking about answers to her questions because I knew she'd answer them right away.</p>
<p><b>Cognitive Engagement</b></p>	<p>HIGH: Had visitors consider whether former inhabitants could have imagined what Yosemite Valley is like today, tied that to having us imagine what it will be like in the future. Had us picture how the valley has changed over time and how strange and foreign it would look to us 100 or 1,000 years from now.</p> <p>HIGH: Walk focused much of the audience's cognitive abilities on imagining what the landscape used to look like and what features used to be there/how they played a role in</p>

Appendix A. Qualitative field notes.

	<p>the battle that took place there. At each stop and walking between them, the ranger regularly reminded visitors to imagine themselves in the places of the soldiers who were there, walking the same lines that they did, and considering the emotions/decisions they faced during the battle. Made the experience really come alive for the audience.</p> <p>HIGH: During the program, the ranger did an exceptional job of portraying the viewpoints of the Confederates, the Union, and civilians in the area of the time. Combining this with a lot of calls to picture or imagine the scenery and the battle at this location really brought the history alive for visitors. He took time to describe what we would have seen if we were sitting with our family having a picnic and watching the battle, or what it would have looked like from the perspective of one of the soldiers.</p> <p>HIGH: Interpreter frequently asked the audience to “imagine” to help them experience the site.</p>
<p><b>Relevance to Audience</b></p>	<p>HIGH: Interpreter made it a priority to connect with and learn a bit about each program participant. He carried on conversations with various visitors between stops, using the information he gathered to shape what he talked about next. He related each story he told to something of particular interest to someone in the audience and kept everyone interested by connecting them in some way to the story being told.</p> <p>HIGH: Compared people coming together in the 1800s after events at Fort McHenry to people coming together after September 11, 2001 and other recent events.</p> <p>HIGH: Described the Civil War as something that took place in back yards and town squares, had us imagine what life would be like now if war broke out in the United States.</p> <p>HIGH: Discussed how global warming can (and already does) affect humans, in terms of flooding, changing weather patterns, and the effect that those things have on our day to day lives, the food sources we rely on, and the economic system that keeps things functioning.</p> <p>HIGH: Ranger's main approach was connecting complex geology to something most people would understand: pizza.</p> <p>HIGH: Allowed visitors to choose some of the locations we visited and had them read out different info with each transition.</p> <p>HIGH: Ranger makes many diverse attempts to create relevance for different people and focuses on looked-over aspects of the Grand Canyon. He told us about the Harvey Girls who worked at the Grand Canyon and how the CCC boys used to flirt with them. He showed us an obviously man-made heart-shaped rock that was laid in the wall facing the girls’ previous dormitories. This was a great connection for me and other young women in the audience; my husband and I later took our pictures with the heart rock. We felt like we had been let in on a Grand Canyon secret that not many other people knew about.</p> <p>HIGH: Ranger created relevance through asking people about activities they used their</p>

Appendix A. Qualitative field notes.

	<p>front porches for and then relating it to the Grant family's porch.</p> <p>LOW: The program provided a ton of factual information about the battle that took place here and the strategies used by either side to gain the upper hand. However, the ranger did not interact with the audience at all and knew nothing about their interests or background. She made no effort to connect the visitors to the resource, either through something of particular interest to them or by creating some relevance between what happened here and the lives of the audience.</p> <p>LOW: The ranger attempted to connect black bears breaking into cars for food to how desperate we would be if we were hungry. If you've ever been starving hungry, you know that you'd be willing to break into a store or steal somebody's lunch....it seems unlikely that anyone on the program has experienced this before or would know what that feels like.</p>
<p><b>Affective Messaging</b></p>	<p>HIGH: The ranger discussed with us the heartache and suffering that went into sending a son off to war or finding out that a loved one had been killed in action. He spoke of the dedication to each other and to country that these soldiers displayed, the determination with which they fought, and the camaraderie on which they relied to keep their spirits up and keep fighting. He showed respect for the memorial by lowering his voice and told us about the importance that their service should have to us. Rather than focusing on numbers or specific dates/battles, he shared the emotional toll that war took on everyone.</p> <p>LOW: This ranger relied solely on historical information to tell the story of FDR and his presidency. He told us the various offices FDR held, explained what polio was, and gave us descriptions of the design/construction of the monument itself. He told us about the impact that war and economic depression had on our country, but only in terms of money and power. He did not include any emotional connection to the struggles of poverty, the despair that people faced, the joy we felt after winning the war, or the emotional toll that polio must have taken on FDR and those around him.</p>
<p><b>Fact-Based Messaging</b></p>	<p>HIGH: This program, about the flora found within the park, provided an abundance of facts and scientific names, but did little to tell us why these plants mattered or what relevance they had to us. The ranger simply listed fact after fact after fact for the duration of an hour long program. After a point, everything began to blend together and lose its meaning.</p>
<p><b>Surprise</b></p>	<p>HIGH: This program was about early western explorers and how their miscalculations/misinterpretations often led them to make the wrong decisions. Throughout the program, the ranger provided us with information that led us to believe we were heading west toward the water, but at the end of the program we came out to a bluff overlooking the water and realized that we were even further away from it than when we had started.</p> <p>HIGH: The ranger turned off the lights in the cave to illustrate how dark it really is.</p>
<p><b>Novelty</b></p>	<p>HIGH: As a part of this program, we were allowed to enter and explore an exquisitely restored historical home that was otherwise closed to the public. Throughout the program, the ranger referred back to the fact that this was an incredibly unique and valuable place, and that we were fortunate to be one of few people who got to see it.</p>

Appendix A. Qualitative field notes.

	<p>HIGH: We got to create our own corn meal from resources found within the park, taste it, and experience how early people would have made their own food.</p>
<p><b>Provocation</b></p>	<p>HIGH: The ranger told a very emotional story about how the coast Miwok tribes were torn away from their homes and lifestyle. He reminded us that their descendants are still alive today and that they can no longer visit the historic sites of their families. He reminded us to think about the impact this must have on their culture and pride.</p> <p>HIGH: Interpreter provided an interesting artistic perspective to the program, left time for reflection, and after the talk, gave people the opportunity to perform their own form of art while watching the sunset.</p> <p>LOW: At one point during this program, the ranger mentioned that urban sprawl is slowly taking over habitat and surrounding national parks in different places across the country. This was stated as a fact and then he moved on to the next subject. Rather than digging deeper or encouraging us to think about the effect that this might one day have, he just mentioned it and did nothing more with it.</p>
<p><b>Holistic Story</b></p>	<p>HIGH: This ranger used the unique and sometimes valuable natural resources of the park to illustrate why native people originally settled here, why it inspired people to move westward, how they used these resources to settle and live off the land, how this led to their over-exploitation, and ultimately to their protection. Each stop taught us about a new resource (trees, rock, grazing fodder, minerals, water, etc.) that played a part in this story. As we moved along, so too did the plot of the story being told.</p> <p>HIGH: The ranger had very clear intentions for the program and a strategy to execute this plan. He informed the audience of what we would be talking about and that each stop had some connection to our story. The story progressed linearly through time and each stop represented a new time period. Every stop was tied back to the central theme and was relevant to the story being told. He used the repetition of certain ideas and interactions with the audience to build a story that came to its conclusion at our last stop.</p> <p>LOW: It was a jumble of dry facts about an otherwise interesting animal (bighorn sheep). There were several moments of "Hmm, what else can I tell you..."</p> <p>LOW: During the tour of a historical home, the interpreter listed off different facts and stories as we walked through each room. A piece of furniture or book would cause her to say "Oh, this reminds me about..." None of what she told us seemed to be connected, and although it was interesting, did not tell us a story about the place or why it is worth preserving. In the end, she talked more about which furniture pieces were original or reproductions than about the people who lived in the house we were walking through and their stories.</p> <p>LOW: As we wandered along the path of our guided walk, the ranger pointed out random trees, buildings, or objects. Each one was described in a manner unrelated to the last and we were left wondering what the point of the program was. At times we would sit and talk about a historical figure from the area, then we would stroll on to the next stop looking at things we happened to pass along the way. There was no clear topic or point to the talk and visitors seemed disconnected and bored by the talk.</p>

Appendix A. Qualitative field notes.

<p><b>Intro/ Conclusion Linkage</b></p>	<p>HIGH: This ranger had a purposeful, powerful introduction about the horror and unpredictable nature of war and how it affected everyone. He used various stops throughout the program to illustrate this point, and at the end reminded us that life back then (like now) could change in an instant. As a young man, you could quickly be enlisted in the Army and sent off to fight for your country. As he did so, he pointed out to us that we were standing next to the grave of a young man who had grown up nearby, gone off to war, and been brought back and buried here in his home town.</p> <p>LOW: The ranger went so far past the designated end time of the program that he did not get the chance to wrap it up in any way. Visitors had to leave the program while he was still talking so they could catch the bus back to the VC.</p> <p>LOW: This program was a classic example of a poorly executed conclusion. While it seemed like the ranger was in the middle of his talk, he simply stopped, looked at the audience, and said “ok, well that’s it.” The program ended very abruptly, with no conclusion at all, leaving the audience wondering what the point of the program was. He had all the opportunity in the world to tie things together and leave us with a lasting message to think about.</p>
<p><b>Clear Theme</b></p>	<p>HIGH: This program focused on the power of Yosemite and the influence it has had in so many people’s live throughout time. The ranger described how it had a spiritual power for native people, was a place of unrivaled beauty and reflection for early explorers, and a place of relaxation and escape for people today. Every stop supported the idea that Yosemite is a unique and powerful place worth preserving, which he reinforced by reminding us that future generations have just as much right to experience and gain from this place. He used a clear and powerful theme to tell the audience why Yosemite is worth protecting.</p> <p>LOW: The ranger on this program told us explicitly that he was going to tell us why a historical building was a unique place. We then walked around and through the hall, where he told us that various treaties were signed and historical figures sat. This was the extent of his program. He did not tell us how those documents have shaped our history, what role those figures played in founding our country, or why preserving the building itself should matter to us. The program was a collection of dates and names, but little more.</p>
<p><b>Central Message</b></p>	<p>HIGH: This program focused on climate change and the impact that it can have on our lives. We were told over and over again throughout the program to think about why we should care. No matter what the science or politicians say, the changes that have already occurred are something that will affect us and that we should be thinking about. The ranger used powerful illustrations of flooding, storm damage, and drought to keep us thinking.</p> <p>HIGH: The ranger used powerful emotional language (“the struggle for freedom,” “the ultimate sacrifice,” and “the value of our freedom”) to remind us of why this monument exists and why it should matter to us. He convinced us that it deserves our respect and reverence, not because of what the monument is itself, but because of who it represents.</p> <p>LOW: During the course of this program, the ranger talked about boats, earthquakes, sea</p>

Appendix A. Qualitative field notes.

	life, and gold. He was very interesting to listen to and taught the audience a lot of things they likely didn't know before. However, these random topics together did not convey a central message. Rather, it left you with a feeling of "huh, that was interesting," but certainly did not change the way you felt or leave you thinking about any particular take-home message.
<b>Consistency</b>	LOW: Ranger's program seemed oddly split; the first half was a very engaging, tactile program about buffalo, and the second half was an abrupt switch into plant identification. The kids were not so interested in the plants and it was severely hot out on the prairie where the plant part was. The program could have easily just been about the buffalo and it would have been great.

## **APPENDIX B. DETAILED STATISTICAL TESTS**

*Only statistically significant results ( $p \leq 0.05$ ) are shared. Blank cells indicate  $p > 0.05$ .*

**Table B-1. Statistically significant t-tests results, comparing the means of visitor outcome scores for selected categorical variables for programs with five or more attendees.**

Observed category	Satisfaction			Visitor experience and appreciation			Behavioral change		
	Mean difference	t-statistic	p-value	Mean difference	t-statistic	p-value	Mean difference	t-statistic	p-value
Impatience	-0.36	-2.2	.031	-0.47	-3.3	.001			
“Friend” identity	0.23	2.3	.023						
“Walking encyclopedia” identity							-0.20	-2.2	.031
Fact-based messaging	-0.34	-3.9	< .001	-0.12	-2.6	.011			
Unexpected negative circumstance	-0.29	-2.8	.006	-0.19	-3.6	< .001			

The following categorical variables yielded no statistically significant differences in visitor outcomes: Inequitable treatment of the audience, questionable information, “Authority” identity, unexpected positive circumstances, use of props.

**Table B-2. One-way ANOVA comparing outcome variables for programs of different pace with five or more attendees. Items not sharing the same superscript are statistically different from one another.**

Pace	Means		
	Satisfaction	Visitor experience and appreciation	Behavioral change
Too fast	8.62 <sup>A</sup>	4.27 <sup>AB</sup>	2.56 <sup>A</sup>
Too slow	8.43 <sup>A</sup>	4.23 <sup>A</sup>	2.84 <sup>AB</sup>
No problems	9.03 <sup>B</sup>	4.44 <sup>B</sup>	2.96 <sup>B</sup>
Statistics	F = 12.9; p < .001	F = 6.9, p = .001	F = 3.2, p = 0.42

**Table B-3. Statistically significant t-tests results, comparing the means of visitor outcome scores for interpreters who expressed different intended outcomes for their interpretive programs.**

Intended outcome	Satisfaction			Visitor experience and appreciation			Behavioral change		
	Mean difference	t-statistic	p-value	Mean difference	t-statistic	p-value	Mean difference	t-statistic	p-value
Increased knowledge				-0.12	2.4	0.019			
Increase desire to learn	0.20	2.2	0.029	0.14	3.2	0.002			
Change attitude	0.18	2.0	0.048	0.16	4.3	< 0.001			
Increase appreciation for Park	0.22	2.7	0.007	0.09	2.2	0.028			
Increase understanding of resource				0.08	2.1	0.040			
Increase level of concern							0.27	2.2	0.032
Change visitor behavior							0.41	2.7	0.008

## **APPENDIX C. CONTEXTUAL COMPARISONS**

To examine whether different predictors of outcomes were more or less prominent in certain situations, we split the sample based on contextual variables and re-examined relationships between predictors and outcomes within each subsample. For each analysis we used the sample of programs with more than five attendees. We looked for meaningful differences in correlation coefficients between predictor variables and outcomes and their statistical significance across the subsamples. Where differences were noted, we explored the distributions of each variable to determine whether the difference could be attributed to a lack of variability within a subsample or a potentially meaningful difference between the subsamples. We share summary tables first (Tables C-1 and C-2) and more detailed statistical tables at the end of the Appendix (Tables C -3, C-4, C-5, and C-6).

While the results show some potentially meaningful differences in the effectiveness of particular program characteristics across contexts, ***these differences are speculative***. While we are confident that our overall sample represents a reasonable approximation of the diversity of interpretive programs across the NPS, we can be less confident that each subsample adequately represents all urban, remote, natural, or cultural programs individually. As the sample size is reduced, generalizability is weakened. As such, ***we share the results of contextual analyses as hypotheses that could be further researched to test their validity***.

We first examined the data within the context of different park locations: urban, proximate, and remote. Within our sample of programs with five or more attendees, 91 programs took place in urban parks, 50 took place in proximate parks, and 131 took place in remote parks. There were no significant differences in outcomes based upon park locations. However, certain variables were more predictive of outcomes in certain areas. The small number of programs within the proximate park subsample inhibited clear interpretation of differences between proximate and other park units. We thus explore only differences between urban and remote park units.

Table C-1 summarizes relationships between predictor variables and outcomes in both urban and remote parks. Only predictor variables showing at least one statistically significant relationship with an outcome are shared.<sup>4</sup> Pluses and minuses indicate statistically significant relationships and their direction. Double pluses indicate significance at  $p < 0.01$ , and single pluses and minus indicate significance at  $p < 0.05$ . Shaded gray cells indicate no statistically significant relationship. For example, *confidence* and *authentic emotion and charisma* were statistically significant positive predictors of *behavioral change* in remote park units ( $r = .265$ ,  $p = .003$  and  $r = .262$ ,  $p = .003$ , respectively), but not in urban park units.

These analyses reveal some nuances within the broader results. First, while the influence of the quality of humor was consistent across park types, *humor quantity* and *sarcasm* were linked to more positive attitudinal outcomes (*satisfaction* and *visitor experience and appreciation*) only in urban park units. *Multisensory engagement* and *novelty* were positively linked with attitudinal outcomes at urban park units, but not remote ones, and *formality* showed a negative relationship with these outcomes only at urban parks. While programs that were shorter than planned were negatively correlated with these outcomes at both types of park units, programs that ran longer were only linked to more positive attitudes at urban parks.

---

<sup>4</sup> In some cases, low variability within one subsample inhibited the interpretation of statistical tests within that subsample.

Appendix C. Contextual comparisons

A number of characteristics were statistically linked to behavioral outcomes at remote parks only. These included variables showing negative relationships with behavioral change, including *sarcasm* and *false assumptions about the audience*. Meanwhile, *confidence*, *authentic emotion and charisma*, *appropriateness for the audience*, *multisensory engagement*, the presence of a *clear message*, and the *quality of the resource* were positively linked to behavioral change at remote park units, but not urban ones.

**Table C-1. Differences in the power of predictive variables in urban and remote parks.**

Characteristic	Satisfaction		Visitor experience and appreciation		Behavioral intention	
	Remote	Urban	Remote	Urban	Remote	Urban
Confidence index	++	++	++	+	++	
Authentic emotion and charisma index	++	++	++	++	++	
Appropriate for the audience	++	++	++	++	++	
Connection	++	++	++	+		
False assumption about audience	--		--		-	
Consistency	++	++	++	+		
Organization	++	++	++	++		
Responsiveness	++	++	+	++		
Humor quality	++	++	+	+	+	
Clear message	++	+			++	
Surprise	+		+		++	
Verbal engagement	+	++	+	++		
Resource quality	+		++			
Appropriate logistics			++	++		+
Audibility		+		+		
Formality				-		
Character acting						+
Novelty		++				
Multisensory engagement		++			+	
Sarcasm		++		+	-	
Humor quantity		++		++		
Fact-based messaging	--	--		-		
Program at least 20% longer than advertised		++		+		
Program at least 20% shorter than advertised	--	--				
Appropriate pace	++	+	++			

We ran a set of similar analyses for program types: nature-focused vs. culture or history-focused (Table C-2). We again limited our analyses to programs with five or more attendees. For this analysis, we removed programs with equally balanced nature-based and culturally-based content because of their

Appendix C. Contextual comparisons

small sample size (n = 29). There were 70 nature-focused programs and 170 cultural/history-focused programs with five or more attendees within the sample for analysis. *Behavioral change* scores were statistically higher for nature-based programs (means: 3.05 vs. 2.84, t = 2.2, p = 0.026). No other statistically significant differences were noted in overall outcomes. Interpreters were more likely to express *behavioral change* as an intended outcome for nature-focused programs as opposed to culturally-focused programs ( $\chi^2 = 7.4$ ; p = .007).

Overall, a number of behavioral predictors were stronger for nature-focused programs, including *confidence, authentic emotion and charisma, appropriateness for the audience, fact-based messaging* (negatively related), and *surprise*. *Humor, audibility, and sarcasm* (negatively related) were better predictors of attitudinal outcomes for cultural programs.

**Table C-2. Differences in the power of predictive variables for natural vs. cultural programs**

Characteristic	Satisfaction		Visitor experience and appreciation		Behavioral intention	
	Natural	Cultural	Natural	Cultural	Natural	Cultural
Confidence	++	++	++	+	++	
Appropriate for the audience	++	++	++	++	+	
Authentic emotion and charisma	++	++	+	++	+	
False assumption about audience	--		-			
Connection	++	++	++	++		
Clear message	++	++		++		
Consistency	+	++	++	++		
Verbal engagement	+	++	++	+	+	
Appropriate logistics	+			++	+	+
Multisensory engagement	+	++	+			
Organization	+	++	+	++		
Novelty	+					
Responsiveness		+	++	+		
Humor quality		++		++		
Surprise					+	
Audibility		++		+		
Humor quantity		++		++		
Sarcasm					--	
Fact-based messaging	-	-				
Program at least 20% shorter than advertised		--				
Appropriate pace	+	++	+	+		+
Use of props			+			

**Table C-3. Correlation coefficients for urban vs. remote comparisons.**

Characteristic	Satisfaction		Visitor experience and appreciation		Behavioral intention	
	Remote	Urban	Remote	Urban	Remote	Urban
Confidence index	.519**	.453**	.294**	.264*	.265**	.191
Sincerity index	.432**	.415**	.280**	.352**	.262**	.069
Appropriate for the audience	.366**	.371**	.344**	.391**	.233**	.165
Connection	.364**	.394**	.285**	.270*	.154	.080
False assumption about audience	-.308**	-.096	-.259**	-.189	-.176*	-.039
Consistency	.307**	.466**	.245**	.239*	.148	.178
Organization	.300**	.385**	.353**	.347**	.022	.095
Responsiveness	.275**	.373**	.207*	.355**	.141	.198
Humor quality	.250**	.285**	.201*	.267*	.202*	.107
Clear message	-.226**	-.260*	-.098	-.159	-.297**	-.015
Surprise	.197*	.109	.190*	.068	.278**	-.150
Verbal engagement	.190*	.285**	.199*	.279**	.147	.047
Resource quality	.184*	.070	.229**	-.008	-.008	.187
Appropriate logistics	.162	.186	.240**	.307**	.167	.233*
Audibility	.159	.238*	.043	.267*	.000	.163
Formality	-.132	-.046	-.086	-.259*	-.039	.100
Character acting	.094	.131	.063	.098	-.008	.211*
Novelty	.084	.276**	-.082	.127	-.077	-.025
Physical engagement	.081	.084	.154	.069	.045	.018
Multisensory engagement	.076	.316**	.066	.076	.194*	.047
Personal sharing	.060	-.027	.044	.073	.107	-.024
Sarcasm	.007	.290**	-.070	.259*	-.214*	.051
Humor quantity	-.019	.355**	-.061	.372**	-.027	.163

\*\* Significant at  $p \leq 0.01$ \* Significant at  $p \leq 0.05$ **Table C-4. T-tests for urban vs. remote programs.**

Program characteristics	Satisfaction				Visitor experience and appreciation				Behavioral change			
	Urban		Remote		Urban		Remote		Urban		Remote	
	Mean diff.	t	Mean diff.	t	Mean diff.	t	Mean diff.	t	Mean diff.	t	Mean diff.	t
Fact-based messaging	-0.57	-3.5**	-0.35	-3.0**	-0.23	-2.5*	-0.10	-1.5	-0.06	-0.4	-0.21	-1.8
Program at least 20% longer than advertised	0.51	2.8**	0.06	0.5	0.18	2.0*	0.05	0.6	0.19	1.0	0.04	0.3
Program at least 20% shorter than advertised	0.71	-3.4**	-0.38	-2.6**	-0.13	-1.2	-0.09	-1.1	-0.32	-1.5	0.0	0.1
Appropriate pace	0.46	2.2*	0.43	3.4**	0.19	1.8	0.23	3.2**	0.39	1.9	0.14	1.1

\*\* Significant at  $p \leq 0.01$ \* Significant at  $p \leq 0.05$

**Table C-5. Correlation coefficients for natural vs. cultural programs.**

Characteristic	Satisfaction		Visitor experience and appreciation		Behavioral intention	
	Natural	Cultural	Natural	Cultural	Natural	Cultural
Confidence	.503**	.437**	.297*	.270**	.330**	.112
Appropriate for the audience	.458**	.355**	.492**	.351**	.269*	.122
Authentic emotion and charisma	.440**	.394**	.294*	.316**	.291*	.070
False assumption about audience	-.368**	-.040	-.273*	-.133	-.206	-.041
Connection	.335**	.360**	.311**	.288**	.215	.090
Clear message	.310**	.243**	.212	.201**	.186	.128
Consistency	.302*	.271**	.319**	.253**	.131	.045
Verbal engagement	.290*	.212**	.457**	.177*	.247*	.089
Appropriate logistics	.286*	.115	.222	.247**	.252*	.156*
Multisensory engagement	.282*	.244**	.245*	.109	.183	.031
Organization	.266*	.431**	.276*	.247**	.190	.128
Novelty	.261*	.111	.147	-.069	-.029	-.009
Responsiveness	.207	.208*	.319**	.213*	.035	.015
Humor quality	.202	.277**	.150	.248**	.204	.131
Surprise	.174	.130	.161	.134	.261*	.041
Audibility	.029	.221**	.014	.190*	.056	.120
Humor quantity	-.024	.217**	-.093	.198**	-.033	.039
Sarcasm	-.068	.128	-.083	.074	-.322**	-.049

\*\* Significant at  $p \leq 0.01$

\* Significant at  $p \leq 0.05$

**Table C-6. T-tests for cultural vs. natural programs.**

Program characteristics	Satisfaction				Visitor experience and appreciation				Behavioral change			
	Cultural		Natural		Cultural		Natural		Cultural		Natural	
	Mean diff.	t	Mean diff.	t	Mean diff.	t	Mean diff.	t	Mean diff.	t	Mean diff.	t
Fact-based messaging	-0.34	-2.6*	-0.31	-2.1*	-0.11	-1.9	-0.11	-1.3	0.01	0.1	-0.30	-1.9
Program at least 20% shorter than advertised	-0.53	-3.6**	-0.28	-1.5	-0.10	-1.1	-0.04	-0.4	-0.23	-1.5	-0.01	0.0
Appropriate pace	0.52	3.8**	0.46	2.4*	0.17	2.5*	0.11	2.2*	0.29	2.1*	0.11	0.5
Use of props	0.07	0.5	0.13	1.0	0.01	0.1	0.17	2.2*	0.02	0.2	-0.04	-0.2

\*\* Significant at  $p \leq 0.01$

\* Significant at  $p \leq 0.05$

## Marc J. Stern



**Position:** Associate Professor, Department of Forest Resources and Environmental Conservation, College of Natural Resources and Environment, Virginia Tech.

**Email:** [mjstern@vt.edu](mailto:mjstern@vt.edu)

**Office Phone:** 540-231-7418

**Education:**

B.S. (Natural Resources), Cornell University, Ithaca, NY  
M.E.Sc. (Environmental Science), Yale University, School of Forestry and Environmental Studies  
Ph.D. (Social Ecology), Yale University

**Activities and interests:**

Marc J. Stern is an Associate Professor in the Department of Forest Resources and Environmental Conservation at Virginia Tech. His research focuses on the human dimensions of protected areas management, natural resource planning, community-based conservation, environmental policy, and environmental education and interpretation. Recent studies include research on the relationships between national parks and their neighboring communities, in the United States, the Caribbean, and Southern Ecuador (including Great Smoky Mountains National Park and U.S. Virgin Islands National Park), on the institutional resilience of community-based conservation in Nepal, on Forest Service planning processes associated with the National Environmental Policy Act, on the effectiveness of community-based conservation projects in the United States, on multi-party collaboration in landscape level restoration initiatives, and on evaluating environmental education and interpretation initiatives at multiple sites within the National Park Service and beyond. Marc's primary interests lie in people's responses to environmental projects, management, and communications. His work often focuses on the relationships within and between natural resource agencies, conservation groups, and numerous facets of the public. At Virginia Tech, he teaches courses in Social Science Research Methods, Environmental Education, Environmental Interpretation, Human Dimensions of Natural Resource Management, and Conflict Resolution.

## Robert B. Powell



**Position:** Associate Professor, Department of Parks, Recreation, and Tourism Management and School of Agriculture, Forest, and Environmental Sciences, Clemson University. Director of Research, Institute for Parks, Clemson University.

**Email:** [rbp@clemson.edu](mailto:rbp@clemson.edu)

**Office Phone:** 864-656-0787

**Education:**

B.A. (Interdisciplinary Studies), University of North Carolina, Chapel Hill

M.E.M. (Environmental Management), Yale University, School of Forestry and Environmental Studies

Ph.D. (Forestry and Environmental Studies), Yale University

**Activities and interests:**

Bob Powell is an Associate Professor in the Department of Parks, Recreation, and Tourism Management and in the School of Agricultural, Forest, and Environmental Sciences at Clemson University. His research and outreach program focuses on environmental education and interpretation, international protected area management and ecotourism, and integrated landscape scale planning. As part of his research program, Bob is currently investigating the effectiveness of community-based conservation projects for the National Audubon Society and the Toyota Foundation; developing a service wide assessment tool for Interpretation, Education, and Visitor Orientation for the National Park Service; and evaluating education and outreach in Everglades and Great Smoky Mountains National Parks. In addition he is part of an interdisciplinary team of researchers that is investigating the relationship between parks, human health, and ecological structure using supercomputing available at Clemson University. Past projects include conducting an integrated marine protected area, tourism, and coastal zone planning project in Sri Lanka for US Agency for International Development; working with land managers and tour operators in Antarctica, the Galapagos Islands, and the Colorado River through Grand Canyon National Park to evaluate nature-based tourism as a conservation and informal environmental education tool; and evaluating the Leave No Trace and TreadLightly! outdoor ethics education programs occurring on U.S. public lands.