

Technical Report for the  
United States National Park Service:

GRAND TETON NATIONAL PARK  
INTERPRETIVE PROGRAM STUDY  
2015-2016

Dr. Pat Stephens Williams and Dr. Ray Darville  
Arthur Temple College of Forestry and Agriculture  
Stephen F. Austin State University

Last updated: May 23, 2017

## Executive Summary

As part of finding their path for the next hundred years, the National Park Service is exploring diverse ways to engage the public and help create systemic changes in the way that the public interacts with each other. Facilitated Dialogue in interpretive programs is one of those ways. Traditionally, programming has been based on the expert and delivery, whereas the new direction leans toward an audience-centered, facilitated experience. To determine how the public is responding to this shift, Grand Teton National Park (GRTE) conducted studies in 2015 and 2016. This study included multiple approaches to determine how Facilitated Dialogue was being used, its effectiveness in achieving the goal of increased audience-centered participation, and how visitor participants perceived their experiences in facilitated dialogue and traditional programs. Data collection was conducted through observation and video of programs, examination of program and visitor characteristics of 56 live interpretive programs, and an online, at-home survey for program participants. It is important to note that at no time were the interpreters being assessed on their personal abilities, though several did approach us about wanting to compare their videos over two years for personal assessment and self-coaching. This study was focused on the technique/tool itself.

Our goals were to compare traditional programs with facilitated dialogue programs and to compare program and visitor characteristics and perceptions over these two years.

### **Selected Results:**

- Findings indicate that traditional programs were significantly more attended than facilitated dialogue programs. However, when examining program characteristics, facilitated dialogue programs received significantly higher program evaluation

scores than traditional programs. Adherence to the four-step ARC of Dialogue model was strongly and positively correlated with program characteristics.

- Programs in all four districts were studied with visual and audio recording taking place at all selected programs.
- In 2015, the total number of attendees recorded at the programs was 747 with about 29 (M = 28.5) individuals attending each program. In 2016, we observed 1,147 attendees (M = 38.2). In both years, there were more males (M = 12.3 in 2015 and 17.6 in 2016) than females (M = 16.4 in 2015 and 20.6 in 2016).
- Of the 26 programs (Table 3), 42% (n = 11) were studied at Colter Bay, 23% (n = 6) at Moose, 19% were studied at Jenny Lake (n = 5), and 15% (n = 4) were studied at LSR Preserve.
- Some 98.5% persisted in traditional programs while 90.5% persisted in the facilitated dialogue programs. The mean difference of 1.1% was not statistically significant (t = 0.331, df = 54, n.s.).
- Facilitated programs were rated significantly higher on cognitive engagement (mean difference = 1.54). In addition, facilitated dialogue programs were rated significantly higher in relevance to audience, affective messaging, and provocation while traditional programs were rated higher in fact-based messaging. Finally, facilitated programs were rated significantly higher in verbal engagement and multiple activities.
- We measured whether or not each a program included the entire ARC (4 phases). The results (Table 11) show that 12 of the 15 (80%) traditional programs not reach the first step of the ARC, but one traditional program did complete the entire ARC. And, among the facilitated dialogue program, 10 of the 11 programs reached at least step 3 of the 4 while 1 program did not reach any of the ARC steps.
- Traditional programs, on average, had more attendees, more males, and more whites compared to facilitated dialogue programs. Twenty-seven characteristics were studied using an ordinal classification system for each program (see Appendix A). Programs were delivered professionally by competent interpreters.
- In terms of ARC adherence, the eleven facilitated dialogue programs averaged 3.4 of the 4 steps. Three programs made it to step 3 while 7 of the programs made it all the way through step 4. There were significant mean differences between facilitated dialogue and traditional programs, suggesting that interpreters were approaching and delivering programs appropriately based on program type.
- Just over half (52.4%), reported that this visit was their first visit to Grand Teton National Park. For all respondents mean number of visits was just over 1 (M = 1.42, SD = 3.25), but among only those who had visited the Park prior to this summer the mean was 3.14 (SD = 4.25). Only about 10% of respondents had visited the Park in three or more times and for one respondent in seven this was only their second park visit.

- Four factors (time, location, day, and content) all had mean score above 4.0 with time receiving the highest mean. Program title (M = 3.52) and program format (M = 3.50) were considerably lower than the first four factors in importance.
- Twelve percent of respondents attended more facilitated dialogue programs than traditional programs, 47% attended the same number of each, and 41% attended more traditional programs.
- Over 56% said they preferred the traditional program while 44% said they preferred the conversational program type.
- As a follow-up question, we asked respondents to indicate the amount of audience conversation they prefer during programs. Answers ranged from very little to a great deal (Table 20). Only one-in-seven said they prefer very little audience conversation and about the same percent said, “quite a bit” or “a great deal.” Median response was “some.”

### **Recommendations Based on Data:**

- GRTE staff should continue both traditional and facilitated dialogue programs. It's clear these two program types both have value and impact.
- GRTE staff should continue to conduct an attendee response study. This study should not only measure participant perceptions of and opinions toward both program types, but analysis should be conducted to better determine program type strengths and weaknesses. In addition, if GRTE and NPS truly want to impact behavior and attitude change, a more longitudinal study is needed to confirm or dispel changes.
- GRTE should consider the fact that the two program types may be drawing different population types. Whether or not GRTE desires different participant types is a question that should be addressed.
- GRTE staff should consider the fact that the four districts did not offer facilitated dialogue programs in the same relative proportions. Jenny Lake had the greatest proportion of facilitated dialogue programs while Colter Bay exhibited the smallest proportion of facilitated dialogue programs. Should there be more uniformity? This is merely a question to be considered based on the character of the district and its visitors.
- Attendee persistence to program's end may be a problem to be addressed. Almost 10% of original program attendees left before the end of programs. The percent was slightly, but not significantly, higher for traditional versus facilitated dialogue programs. This percentage also varied somewhat by advertised length of program with proportionally more attendees leaving for longer programs.
- GRTE should evaluate marketing of facilitated dialogue programs, which should include names, topics, locations, and topics. Three of the facilitated dialogue programs did not “make.” This represents about 20% of facilitated dialogue programs studied. It may be necessary to reevaluate which programs, taking into consideration all variables such as time, content, etc., work better as facilitated, traditional, or hybrid.

- Also in the marketing strain, GRTE may wish to reevaluate the titles of the programs to consider giving them different titles that may be more descriptive and appealing to potential audiences.
- Develop more family activities, which was a theme emerging for the survey. Families want to visit the park and develop memories. These could include more ranger-led hikes, fishing, backpacking, and other activities.
- Consider altering park programs based on these results. The keys to park selections were, in large part, practical reasons such as time of day. Some park programs are poorly attended. Consider eliminating those programs or reducing the number of offerings so staff can be deployed for other programs or in different ways. Consider having more public meetings with park visitors for senior staff to discuss park issues, challenges, and solutions.
- Continue to provide opportunities for student participation in research and interpretation. Students participating in this project have truly had transformative experiences which will make them better researchers and professionals in the field.
- The methodology used for this multi-faceted study is an appropriate template to share and use at other parks. Any one of the pieces may be broken out and utilized, however, the observations based on Stern's work, the videography, the visitor survey, and MBTI assessment of the interpreters all added to help build a more complete picture of facilitate dialogue integration and how it may continue to grow as process to not only benefit the park, but also the visitor and their home communities.
- Given the potential of the experience of facilitated dialogue to change its participants, we suggest that GRTE consider continued research in this area on the changes it may instigate in the interpreters themselves. Most of the interpreters are seasonal so it would be interesting to find out how these skills are being used outside of the park. Since one of the reasons behind incorporating facilitated dialogue was in hope of providing the participant with experience in skills that would transcend the park visit and carry over into community action, the interpreters would be a natural place from which to start capturing that proof.
- Conduct more visitor research. Parks are people and society and both change over time.

## Table of Contents

<b>Executive Summary</b> .....	2
<b>Selected Results:</b> .....	2
<b>Recommendations Based on Data:</b> .....	4
<b>List of Tables</b> .....	8
<b>Introduction</b> .....	9
<b>Literature Review</b> .....	11
<b>Facilitated Dialogue as an Interpretive Technique</b> .....	11
<b>Interpretation Research</b> .....	12
<b>Research Methodology</b> .....	14
<b>General Methodology</b> .....	14
<b>Site Description</b> .....	14
<b>Programs Attended</b> .....	17
<b>Data Collection</b> .....	17
<b>Pilot Test</b> .....	19
<b>Data</b> .....	19
<b>Results – Part I: On-Site Program Analysis</b> .....	20
<b>Program Attendees</b> .....	20
<b>General Program Characteristics</b> .....	22
<b>Interpretative Program Characteristics</b> .....	30
<b>Program Characteristics</b> .....	34
<b>ARC Achievement</b> .....	36
<b>By District</b> .....	37
<b>Relationship between ARC Total and Program Characteristics</b> .....	39
<b>Conclusions: Part I</b> .....	40
<b>Results: Part II – Participant Survey</b> .....	43
<b>Demographic Characteristics</b> .....	43
<b>GRTE Visits</b> .....	44
<b>Opinions Concerning the Park</b> .....	44
<b>Programs Attended</b> .....	46
<b>Impact of Programs Attended</b> .....	51
<b>Conclusions: Part II</b> .....	55

**References** .....60  
**Appendix A** .....61  
**Appendix B** .....66

## List of Tables

Table 1: Demographic Characteristics of Program Attendees (n = 59).....	21
Table 2: Demographic Characteristics of Program Attendees by Program Type (n =23).....	22
Table 3: District of Program .....	22
Table 4: Day of Program .....	23
Table 5: Program Start Time .....	24
Table 6: Program Type .....	24
Table 7: Advertised Program Length.....	25
Table 8: Actual Program Length.....	26
Table 9: Program Length Difference .....	28
Table 10: Descriptive Statistics of Program Characteristics (2015) .....	30
Table 11: Descriptive Statistics of Program Characteristics (2016) .....	31
Table 12: Descriptive Program Characteristics by Program Type (n = 26) .....	33
Table 13: ARC Totals.....	36
Table 14: ARC Totals by Program Type.....	36
Table 15: ARC Total by District .....	38
Table 16: ARC Total by Time of Day Program .....	38
Table 17: Length of Program affecting ARC Total :.....	39
Table 18: Descriptive Statistics of Opinions Toward the Park .....	45
Table 19: Importance of Selected Factors in Choosing Programs.....	46
Table 20: Amount of Audience Conversation Preferred .....	49
Table 21: Participate in the Junior Ranger Program.....	49
Table 22: Program Attendance in Descending Order .....	50
Table 23: Factors Influencing Attitude or Behavioral Outcome in Descending Order ALL Programs .....	52
Table 24: Factors Influencing Attitude or Behavioral Outcome in Descending Order Participants Attending at Least One Facilitated Dialogue Program .....	52
Table 25: Comparison of Traditional Programs with Facilitated Dialogue Programs on Influence of Attitude and Behavior .....	54

## Introduction

Historically, park interpreters have been essentially “sages on stages” as they have dispensed facts and figures in their interpretative programs about their respective parks. Anyone having visited Grand Canyon, Grand Teton, Yellowstone, or dozens other national parks can envision the ranger dressed in uniform standing and providing what amounts to a lecture, or talk, to visitors. This essentially one-way communication has been used for decades and is a time-honored tradition. Many individuals visiting parks seek out these experiences and enjoy them greatly. However, as the National Park Service prepared to celebrate its centennial on August 25<sup>th</sup>, 2016, some park managers and interpreters considered new approaches to personal interpretation in order to more effectively reach park visitors and focus on a more audience-centered communication technique. While they have maintained this interpretative approach historically, the utilization of a facilitated dialogue approach is being integrated into the traditional programming. This approach is a two-way form of communication designed to more actively engage visitors in critical thinking and communication about park values, goals, and meanings, as well as, hopefully, help provide skill development for its participants to be able to take back and use in their own communities. This approach is more of a “guide on the side” approach in which visitors are encouraged to talk with the ranger and with each other in a focused program.

This research study of interpretive programs in Grand Teton National Park had three broad goals: (1) to examine facilitated dialogue and traditional interpretative programs and their characteristics, (2) to study attendees and their characteristics and perceptions and preferences on facilitated dialogue programming, and (3) to pilot test a

method to measure facilitated dialogue programs. In particular, the study's focus was on analyzing the facilitated and traditional programs offered during Summers 2015 and 2016 and audience response to those programs. Guided by park administration, we sought to answer these research questions through the research:

1. How many individuals attended selected facilitated and traditional interpretative programs in Grand Teton National Park during summer, 2015?
2. What were the observable characteristics of these individuals?
3. Do the characteristics of those attending facilitated dialogue programs differ from those attending traditional interpretative programs?
4. What were the observable characteristics of these programs?
5. What were the responses of attendees to these programs?

Our specific objectives for this study are:

1. Determine how to measure Facilitated Dialogue effectiveness;
2. Determine visitor response to Facilitated Dialogue in comparison to traditional programming;
3. Design and test a template for this research that may be adapted in other parks;
4. Collect examples of Facilitated Dialogue in Action for future training purposes;
5. Reassess current coaching process to determine if changes need to be made;
6. Determine best practices for Facilitated Dialogue in the parks;
7. Determine if MBTI (Myers-Briggs Type Indicator) can predict those rangers more inclined to use the Facilitated Dialogue method over the Traditional or Hybrid methods;
8. Determine if content and context guides the choice of Traditional, Hybrid, or Facilitated Dialogue for a program.

## Literature Review

### Facilitated Dialogue as an Interpretive Technique

Recently, professionals have sought to create a more active, engaged interpretative experience for visitors to parks, museums, and other appropriate venues. Facilitated dialogue (National Park Service, 2014) is a “form of interpretative facilitation that uses a strategically designed set of questions – an “arc of dialogue” – to guide participants into a structured, meaningful, audience-centered conversation about a challenging or controversial topic.”

This interpretative technique is designed to assist visitors to have a structured conversation centered on potentially controversial topics such as climate change, race/ethnic relations, slavery, endangered species, and other such areas of interest. Interpreters employing this approach are encouraged to move the participants through four consecutive phases: (1) Building Community, (2) Sharing Personal Experiences, (3) Exploring Beyond Our Own Experience, and (4) Synthesizing and Bringing Closure to the Dialogue. In “Building Community,” the interpreter (facilitator) builds safety and acceptance and provides information regarding the topic and the purpose of the dialogue. During this phase, a facilitator should engage in a variety of techniques to “break the ice” and to have participants develop a sense of community and safety in this non-traditional type of program (United States National Park Service, 2014). A number of these techniques may be found elsewhere. Sample questions might be to ask attendees where they are from, or to ask them to describe themselves in five words or less. At the end of this phase, attendees should feel welcomed and not threatened. They should feel

accepted. They should be ready for more personal sharing when, in phase 2, they are asked to share a meaningful, personal experience.

In phase 2, sharing our experiences, a facilitator asks participants to think about their own experiences concerning the topic and to share those experiences with one another. Questions to generate this form of sharing should be open-ended and welcoming so that each individual is perceived to be equal among group members. This sharing should be done non-judgmentally. Questions should allow, even encourage, members to share different experiences, not simply similar experiences. This phase allows attendees to begin to think about other's experiences, which are different from their own.

Then, in phase 3, exploring experiences of others, a facilitator encourages participants to move beyond their own experiences so that they engage in inquiry and learn with and from one another (especially those whose experiences are different from their own).

Finally, synthesizing and bringing closure to the dialogue (phase 4), a facilitator helps participants to connect ideas, perspectives, and information gained during the program. Ideally, the facilitator will use this model (or something similar) as he or she plans the program. And, ideally, the facilitator makes it all the way through all four phases of the arc so that there is a sense of closure to the program.

### **Interpretation Research**

The portion of our research here that is trying to determine an appropriate way to measure facilitated dialogue is based on a recently-conducted national study of live interpretive programs in the United States Park Service (Stern, et.al, 2012), sponsored

by the National Park Conservation Association's Center for Park Management. In 2011, a team of four researchers attended 376 live interpretive programs in 24 NPS cultural and natural resource units. Researchers conducted brief interviews with interpreters and collected data on 56 characteristics associated with the programs and their interpreters. And, they conducted post-program surveys among the visitors, collecting data from some 3,603 program attendees. These researchers measured program outcomes, including level of satisfaction with the program, visitor experience and appreciation, and behavioral change among visitors. Their work was designed to see if they could identify best practices, which were associated with positive program outcomes. Fifteen best practices were identified, ranging from confidence to avoiding making uncertain assumptions about the audience.

## **Research Methodology**

### **General Methodology**

The general methodology for this was unobtrusive observation at selected interpretative programs during Summers 2015 and 2016. Using direct observation with trained research assistants, data collection occurred in Grand Teton National Park between July 20, 2015 and July 29, 2015 and between July 8, 2016 and July 30, 2016. In 2015, 26 programs were studied in 2015 while 30 programs were studied in 2016, yielding a total of 56 programs. Programs in all four districts were studied with visual and audio recording taking place at all selected programs. In 2016, interpreters announced the opportunity for adult visitors to participate in an online survey. Visitors wishing to participate provided their name and email address. After each program, researchers completed a data collection form, which measured selected program characteristics. Following all programs, a team of two trained researchers evaluated programs based on an established data collection form (Appendix A).

### **Site Description**

Grand Teton National Park is located in western Wyoming, south of Yellowstone and east of the Idaho state line. Long considered a flagship national park, Grand Teton welcomes almost three million visitors annually. The number of recreational visitors is clearly on the increase from a recent low of 1.2 million in 1988 (the year of the Yellowstone National Park fires) to a high of 3.2 million in 1977. Since 2000 mean recreational visitors have averaged 2.6 million, which is about one million above its historical mean of 1.6 million (since 1929). It is annually one of the ten most visited parks

in the United States-8<sup>th</sup> most visited in 2014 (<http://www.npca.org/exploring-our-parks/visitation.html>). In 2015, 3.1 million recreation visitors were recorded, yielding a year-to-year increase of 11 percent from 2014. The number of recreation visitors

The Park, not large in land area compared to others (such as such as adjacent Yellowstone), is comprised of 310,000 acres (480 square miles). While it is not large in land area, it makes up with resources, with beauty, and with its attractions. For example, it includes the 40-mile Teton Range and is named for its highest peak Grand Teton, which rises to about 13,775 feet above sea level and roughly 7,000 feet above the Jackson Valley. The Park was established in 1929 through efforts of many individuals, including John D. Rockefeller, Jr. Many of its visitors return year-after-year and some are now brining their grandchildren. Many visitors go to both Grand Teton and Yellowstone during the same trip, which provides them with different experiences.

The Park is located in the Jackson Hole area and upper Snake River Valley. Some of the rocks in the park have been dated to being 2.6 million years old and are some of the oldest known rocks in the United States. While the glaciers are receding, the mountain peaks inspire awe year-around.

Furthermore, the Park is virtually an intact ecosystem with the same species of flora and fauna that have lived there for thousands of years. There are more than 1,000 known vascular plant species, over 300 species of birds, more than a dozen bird species, as well as dozens of species of mammals including elk, bison, bear, and wolf.

Grand Teton National Park has some unparalleled scenery for those who simply want to take in the sights, but individuals visit the park for other recreational opportunities, including mountaineering, hiking (over 200 miles of trails), fishing, boating,

skiing, and cycling. For those wanting an interpretative experience, the Park maintains visitor centers and ranger stations at Moose (Craig Thomas), Jenny Lake, and Colter Bay (<http://www.nps.gov/grte/planyourvisit/visitorcenters.htm>).

The Craig Thomas Discovery and Visitor Center at Moose is located about 12 miles north of Jackson, Wyoming. It includes an information desk and backcountry desk, restrooms, a large Grand Teton Association bookstore, a mountaineering exhibit, state-of-the-art exhibits, and a park documentary shown in an auditorium. It was designed to showcase the Teton Range, which is in clear view out the back window. This busy visitor center hosts a large number of interpretative programs.

Jenny Lake Visitor Center, located about eight miles north of Moose has geology exhibits and information as well as a relief map of the park. Park rangers distribute park information, answer questions, and give interpretative talks around this busy visitor center.

The Laurance S. Rockefeller Preserve Center is located four miles south of Moose on Moose-Wilson Road. It is the newest of the visitor centers, having been constructed on land donated to the park by the Laurance S. Rockefeller family. This visitor center is unique in its architecture and purpose, providing unique sensory experiences found nowhere else in the park and perhaps in all national parks. It is surrounded by an eight-mile trail network and supports interpretative programs daily.

Located some 25 miles north of Moose, adjacent to Jackson Lake, is the Colter Bay Visitor Center. While an older building, it was remodeled in 2012 and displays a few of the David T. Vernon collection of Indian arts. In the summer, it hosts Native American artists. Included in this very busy visitor center are an information desk, a Grand Teton

Association bookstore, auditorium, restrooms, and a backcountry office. Because of its location and for other reasons, Colter Bay Visitor Center has the largest number of interpretative programs in the park.

### **Programs Attended**

During the period of July 18<sup>th</sup> and July 29<sup>th</sup>, 2015 a total of 29 interpretative programs were attended in each of their Grand Teton National Park four districts: LSR Preserve, Moose, Jenny Lake, and Colter Bay. Of these 29 programs, some program descriptions contained language that supported facilitated dialogue such as *Lakeshore Conversations* at Jenny Lake, while the rest were traditional interpretative programs. Each of the four districts contained at least one facilitated dialogue program. The majority of the programming was conducted in the traditional and/or a hybrid of traditional and facilitated dialogue. In Summer 2016, another 30 programs were selected for study, yielding a total of 59 studied programs. Selections were made based on attempting to balance the selected programs compared to the overall program offerings throughout the Park, type of program, location of program, and logistics (see Appendix X for list of programs). The targeted program list was submitted to Park managers for their review and approved.

### **Data Collection**

Data collection was completed over two summers. Summer 2015 was unobtrusive observation while Summer 2016 included both the observations and an OMB approved visitor survey. In both summers, two researchers attended each studied program. Each

program was recorded using both video and audio capture by using GoPro Hero (1080p30 FPS) devices and Olympus WS-821 Digital Voice Recorders. Researchers participated as attendees, but did not dominate discussion or any other program activities. They remained as unobtrusive as possible. Upon completion of the program, the researchers completed the program demographics form and the program characteristics form (see Appendix A). This data collection form was adapted from the Stern (2012) study, following their form closely. In addition to the characteristics form, research assistants would determine the level of the Arc of Dialogue achieved in the program through a discussion/consensus modeled after the process used in determining the success of IDP programs submitted for certification.

Thirty program characteristics were identified as being particularly important for study objectives. These all related to salient program characteristics and included such program aspects as introduction quality, transitions, verbal engagement, affective messaging, novelty, place-based messaging, and more than 20 others. Measurement conformed to nominal and ordinal level of measurement with some questions employing nominal measurement (such as use of props) with other questions employing ordinal measurement (such as physical engagement with four ordered attributes). A complete list of the variables measured and their attributes can be found in Appendix A. In addition, we collected program-related data:

- name of program
- day of program
- start time of program
- type of program
- location of program
- advertised length of program
- actual length of program
- number of attendees

- number of males
- number of females
- number of adults
- number of children
- number of whites
- number of blacks
- number of Asians

Program video and audio recordings were analyzed for content by certified interpreters. This analysis is presented in another report.

### **Pilot Test**

A pilot test of the data collection techniques was conducted prior to actual data collection. Research assistants attended Lakeshore Conversations at Jenny Lake. This allowed researchers to go through the data collection process, test equipment, and work with the data collection forms prior to actual data collection. After discussion of the data collection process, the data collection form was finalized as was the data collection procedures. No changes were made to either during the data collection phase of the study.

### **Data**

Upon completion of data collection, researchers entered data in an Excel file. An SPSS data file was constructed using standard techniques and then data were transferred into SPSS (Professional Version 23) for analysis. Once in SPSS, randomly-selected data records were checked for accuracy and standard data checking techniques were employed to check for wild codes and coding errors. Finally, standard statistical procedures were conducted to answer research questions and test hypotheses.

## Results – Part I: On-Site Program Analysis

### Program Attendees

We counted the number of individuals, number of males, number of females, number of adults, number of children, number of Whites, number of Blacks, and number of Asians attending each program that was given (Table 1).

In 2015, the total number of attendees recorded at the programs was 747 with about 29 ( $M = 28.5$ ) individuals attending each program. In 2016, we observed 1,147 attendees ( $M = 38.2$ ). In both years, there were more males ( $M = 12.3$  in 2015 and 17.6 in 2016) than females ( $M = 16.4$  in 2015 and 20.6 in 2016). Females outnumbered males in 45 of the programs while males outnumbered females in four programs. Adults were about two-thirds of program attendees while children comprised the other one-third. Whites comprised about 89% of program attendees while Blacks and Asians made up the remaining 11%.

Table 1: Demographic Characteristics of Program Attendees (n = 59)

Demographic Characteristic	Mean	Median	Standard Deviation	Minimum	Maximum	Range	Sum
Total-2015	28.5	19.0	28.7	4	128	124	746
Total-2016	38.2	26.5	30.6	8	142	134	1,147
Males-2015	12.3	8.5	12.8	0	56	56	320
Males-2016	17.6	12.5	14.3	3	62	59	529
Females-2015	16.4	10.0	16.3	3	72	69	427
Females-2016	20.6	15.5	16.3	4	80	76	618
Adults-2015	19.1	11.0	20.6	4	98	94	496
Adults-2016	23.7	18.0	17.7	6	84	78	712
Children-2015	9.6	8.0	9.3	0	40	40	250
Children-2016	14.5	10.0	13.5	0	58	58	435
Whites-2015	24.3	18.0	22.8	4	108	104	631
Whites-2016	33.4	24.5	25.7	6	117	113	1,003
Blacks-2015	0.2	0.0	0.8	0	4	4	5
Blacks-2016	0.4	0.0	1.3	0	6	6	12
Asians-2015	3.3	0.0	6.2	0	25	25	86
Asians-2016	3.6	1.5	5.6	0	30	30	108

Between 2015 and 2016, the number of program attendees in all demographic characteristics increased, which is consistent with the overall increase in summer visitation in the park from 2015 to 2016. Total attendance jumped 10 individuals per program. Independent-samples t tests were performed to determine if there were significant changes to the demographic characteristics of attendees from year to year among the studied programs. Results indicate that none of the demographic characteristics displayed any significant changes from 2015 to 2016. Table 2 shows the variance of demographic characteristics (number of males, females, adults, children, Whites, Blacks and Asians) by type of program (traditional vs. facilitated).

**Table 2: Demographic Characteristics of Program Attendees by Program Type (n =23)**

Demographic Characteristic	Traditional Program Mean	Facilitated Program Mean	t value	p value
Total Attendees	40.3	17.6	3.91	<b>.000</b>
Males	18.3	7.4	4.01	<b>.000</b>
Females	22.1	10.1	3.72	.000
Adults	25.7	11.1	3.85	.000
Children	14.5	6.4	3.24	.002
Whites	34.9	14.9	4.11	<b>.000</b>
Blacks	0.3	0.3	0.23	n.s.
Asians	4.3	1.5	2.37	.022
Percent Whites	88.8	88.4	0.88	n.s.
Percent Males	44.7	38.7	1.95	n.s.
Percent Adults	65.0	75.4	-.09	n.s.

**General Program Characteristics**

Of the 26 programs (Table 3), 42% (n = 11) were studied at Colter Bay, 23% (n = 6) at Moose, 19% were studied at Jenny Lake (n = 5), and 15% (n = 4) were studied at LSR Preserve. Furthermore, we examined the mean number of attendees by district offering the program. Colter Bay had the highest mean per program (M = 35.7, SD = 27.7) while Jenny Lake had the lowest mean per program (M 9.0, SD = 4.5); however, the means were not statistically significant (F = 2.745, p. = n.s.).

**Table 3: District of Program**

District	Frequency	Percent	Valid Percent	Cumulative Percent
Colter Bay	11	42.3	42.3	42.3
Jenny Lake	5	19.2	19.2	61.5
LSR Preserve	4	15.4	15.4	76.9
Moose	6	23.1	23.1	100.0
Total	26	100.0	100.0	

Programs (Table 4) were studied on all seven days of the week with Tuesday (23.1%, n = 6) and Friday (23.1, n = 6) tied for the most. The fewest programs were observed on Monday (3.8%, n = 1). A one-sample test was performed based on equal probabilities (14.3%) of offerings by day. The test was not significant ( $p = n.s.$ ) and thus there is not a statistically significant difference in number of offerings per day.

Table 4: Day of Program

Day of Program	Frequency	Percent	Valid Percent	Cumulative Percent
Monday	1	3.8	3.8	3.8
Tuesday	6	23.1	23.1	26.9
Wednesday	3	11.5	11.5	38.5
Thursday	3	11.5	11.5	50.0
Friday	6	23.1	23.1	73.1
Saturday	4	15.4	15.4	88.5
Sunday	3	11.5	11.5	100.0
Total	26	100.0	100.0	

Over half (53.8%, n = 14) of the studied programs (Table 5) started during mornings, while 34.6% were studied (n = 9) with an afternoon start, and 11.5% (n = 3) were studied with an evening start time. A one-sample test detected statistically significant differences in the program start times ( $p < .0005$ ) suggesting that morning offerings are significantly more likely to be given compared to afternoon and evening programs.

Table 5: Program Start Time

	Frequency	Percent	Valid Percent	Cumulative Percent
Morning	25	44.6	44.6	44.6
Afternoon	25	44.6	44.6	89.3
Evening	6	10.7	10.7	100.0
Total	56	100.0	100.0	

Because one of the study objectives was to compare and contrast traditional and facilitated programs, both types were studied. Of the 56 studied programs, 71% (n = 40) were traditional were studied compared to 28.6% (n = 16) for facilitated dialogue programs. Each district had at least one full Arc program per season. The majority of programs were traditional while some incorporated some techniques of audience-centered engagement.

Table 6: Program Type

	Frequency	Percent	Valid Percent	Cumulative Percent
Traditional	40	71.4	71.4	71.4
Facilitated Dialogue	16	28.6	28.6	100.0
Total	56	100.0	100.0	

Programs of various advertised lengths were studied with six unique program lengths advertised ranging from 20 minutes to 180 minutes. The shortest program was 20 minutes long (7.1%, n =4) while the longest program was 180 minutes long (5.4%, n = 3). Mean program length was 51.4 minutes (SD = 41.6); there was a total of 2,880 minutes (48 hours) of programs studied. Almost 84% of programs were advertised to last 60 minutes or less.

**Table 7: Advertised Program Length**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 20	4	7.1	7.1	7.1
25	1	1.8	1.8	8.9
30	25	44.6	44.6	53.6
45	15	26.8	26.8	80.4
60	2	3.6	3.6	83.9
90	3	5.4	5.4	89.3
120	1	1.8	1.8	91.1
150	2	3.6	3.6	94.6
180	3	5.4	5.4	100.0
Total	56	100.0	100.0	

Actual program length was recorded by the researchers (Table 8). Mean actual time was 47.8 minutes (SD = 36.) with the shortest program lasting only 17 minutes and the longest program lasted 179 minutes. Thirty-eight percent of programs actually lasted 30 minutes or less.

Table 8: Actual Program Length

	Frequency	Percent	Valid Percent	Cumulative Percent
14	1	1.8	1.8	1.8
17	1	1.8	1.8	3.6
19	2	3.6	3.6	7.1
21	1	1.8	1.8	8.9
22	2	3.6	3.6	12.5
23	1	1.8	1.8	14.3
24	1	1.8	1.8	16.1
25	3	5.4	5.4	21.4
26	1	1.8	1.8	23.2
29	2	3.6	3.6	26.8
30	6	10.7	10.7	37.5
31	3	5.4	5.4	42.9
32	1	1.8	1.8	44.6
33	1	1.8	1.8	46.4
34	2	3.6	3.6	50.0
35	3	5.4	5.4	55.4
36	3	5.4	5.4	60.7
37	1	1.8	1.8	62.5
38	1	1.8	1.8	64.3
40	1	1.8	1.8	66.1
45	2	3.6	3.6	69.6
46	1	1.8	1.8	71.4
47	1	1.8	1.8	73.2
52	2	3.6	3.6	76.8
57	1	1.8	1.8	78.6
58	1	1.8	1.8	80.4
67	2	3.6	3.6	83.9
68	1	1.8	1.8	85.7
70	1	1.8	1.8	87.5
85	1	1.8	1.8	89.3
104	1	1.8	1.8	91.1

120	1	1.8	1.8	92.9
131	1	1.8	1.8	94.6
145	1	1.8	1.8	96.4
166	1	1.8	1.8	98.2
179	1	1.8	1.8	100.0
Total	56	100.0	100.0	

Table 9 below shows that some six programs (23.1%) lasted the same time as advertised. Nineteen programs (33.9%) did not go as long as advertised while 8 (14.3%) lasted exactly as advertised while the remaining lasted longer than advertised. The mean difference between advertised and actual was 3.6 minutes and was not statistically significant ( $t = -1.99$ ,  $df = 55$ ,  $p. = n.s.$ ). The results suggest that interpretation programs lasted as long as advertised generally. Certainly many variables affect exact length of a program, such as weather, number of participants, number of children, talkers v non-talkers, questions, etc.

Table 9: Program Length Difference

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid -25	1	1.8	1.8	1.8
-22	1	1.8	1.8	3.6
-11	1	1.8	1.8	5.4
-7	3	5.4	5.4	10.7
-6	1	1.8	1.8	12.5
-5	3	5.4	5.4	17.9
-4	2	3.6	3.6	21.4
-3	1	1.8	1.8	23.2
-2	2	3.6	3.6	26.8
-1	4	7.1	7.1	33.9
0	8	14.3	14.3	48.2
1	4	7.1	7.1	55.4
3	2	3.6	3.6	58.9
5	6	10.7	10.7	69.6
6	1	1.8	1.8	71.4
7	2	3.6	3.6	75.0
8	3	5.4	5.4	80.4
9	2	3.6	3.6	83.9
11	2	3.6	3.6	87.5
14	1	1.8	1.8	89.3
15	1	1.8	1.8	91.1
19	2	3.6	3.6	94.6
22	1	1.8	1.8	96.4
32	1	1.8	1.8	98.2
76	1	1.8	1.8	100.0
Total	56	100.0	100.0	

Finally, we examined persistence of attendees to continue staying at the program until the end of the program. A total of 243 individuals left these 56 programs ( $M = 4.3$ ,  $SD = 9.1$ ). In 21 programs (37.5%), no one left before the end of the program. Overall,

89.9% persisted and stayed throughout the end of the programs. Next, we examined whether the number of attendees persisting through the programs was related to program type. Some 98.5% persisted in traditional programs while 90.5% persisted in the facilitated dialogue programs. The mean difference of 1.1% was not statistically significant ( $t = 0.331$ ,  $df = 54$ , n.s.).

## Interpretative Program Characteristics

Some 22 program characteristics were scored for each program by the researchers based on the pre-developed coding system (Table 10).

Table 10: Descriptive Statistics of Program Characteristics (2015)

Program Characteristic	Mean	S.D.	% of Programs with Each Score					
			1	2	3	4	5	
Organization								
• Introduction Quality (1 – 3)	2.81	0.40	0.0	19.2	80.8			
• Appropriate Sequence (1 – 4)	3.62	0.57	0.0	3.8	30.8	65.4		
• Transitions (1 – 4)	3.73	0.45	0.0	0.0	26.9	73.1		
• Holistic Story (1 – 5)	2.35	1.01	15.4	50.0	26.9	0.0	7.7	
• Conclusion Linked to Introduction (1 – 4)	3.00	0.85	0.0	34.6	30.8	34.6		
• Clear Theme (1 – 4)	3.81	0.40	0.0	0.0	19.2	80.8		
• Appropriate Logistics (1 – 4)	3.73	0.53	0.0	3.8	19.2	76.9		
Connection								
• Connection to Intangibles/Universals (1 – 5)	4.19	0.80	0.0	3.8	11.5	46.2	38.5	
• Cognitive Engagement (1 – 5)	3.54	1.27	3.8	23.1	19.2	23.1	30.8	
• Relevance to Audience (1 – 5)	3.85	1.08	0.0	11.5	30.8	19.2	38.5	
• Affective Messaging (1 – 5)	3.12	1.21	3.8	34.6	26.9	15.4	19.2	
• Fact-Based Messaging (1 – 3)	2.19	0.98	0.0	7.7	15.4	26.9	50.0	
• Place-Based Messaging (1 – 5)	3.85	1.08	3.8	7.7	19.2	38.5	30.8	
• Provocation (1 – 4)	2.42	1.21	30.8	23.1	19.2	26.9		
Audience Comfort (1 – 4)	3.62	0.57	0.0	3.8	30.8	65.5		
Central Message (1 – 4)	3.46	0.58	0.0	3.8	46.2	50.0		
Appropriate to Audience (1 – 5)	3.85	1.08	0.0	0.0	7.7	23.1	69.2	
Multisensory (1 – 3)	1.69	0.79	50.0	30.8	19.2			
Physical Engagement (1 – 4)	2.19	1.13	7.7	15.4	26.9	50.0		
Verbal Engagement (1 – 5)	3.50	1.24	0.0	30.8	19.2	19.2	30.8	
Surprise (1 – 3)	1.62	0.50	38.5	61.5	0.0			
Novelty (1 – 3)	1.92	0.74	30.8	46.2	23.1			
Consistency (1 – 3)	2.96	0.20	0.0	3.8	96.2			
Resource Quality (1 – 3)	2.08	0.80	26.9	38.5	34.6			
Multiple Viewpoints (1 – 3)	2.38	0.52	0.0	19.2	11.5			
Multiple Activities (1 – 4)	2.19	1.76	50.0	7.7	15.4	26.9		
Pace (0 – 2)	2.00	0.00	0.0	3.8	96.2			
Unexpected Negative Circumstance (0 – 1)	0.08	0.27	92.3	7.7				
Unexpected Positive Circumstance (0 – 1)	0.08	0.27	92.3	7.7				

Table 11: Descriptive Statistics of Program Characteristics (2016)

Program Characteristic	Mean	S.D.	% of Programs with Each Score					
			1	2	3	4	5	
Organization								
• Introduction Quality (1 – 3)	2.87	0.35	0.0	13.3	86.7			
• Appropriate Sequence (1 – 4)	3.50	0.68	0.0	10.0	30.0	60.0		
• Transitions (1 – 4)	3.70	0.47	0.0	0.0	30.0	70.0		
• Holistic Story (1 – 5)	2.43	0.94	6.7	60.0	23.3	3.3	6.7	
• Conclusion Linked to Introduction (1 – 4)	3.23	0.73	3.3	6.7	53.3	36.7		
• Clear Theme (1 – 4)	3.85	0.46	0.0	3.3	10.0	86.7		
• Appropriate Logistics (1 – 4)	3.80	0.48	0.0	3.3	13.3	83.3		
Connection								
• Connection to Intangibles/Universals (1 – 5)	4.03	0.72	0.0	0.0	23.3	50.0	26.7	
• Cognitive Engagement (1 – 5)	3.27	1.31	6.7	26.7	26.7	13.3	26.7	
• Relevance to Audience (1 – 5)	4.57	0.73	0.0	6.7	36.7	33.3	23.3	
• Affective Messaging (1 – 5)	2.93	1.23	10.0	33.3	23.3	20.0	13.3	
• Fact-Based Messaging (1 – 3)	2.27	0.52	0.0	3.3	66.7	30.0		
• Place-Based Messaging (1 – 5)	3.87	1.04	-0.0	3.3	40.0	20.0	36.7	
• Provocation (1 – 4)	2.57	1.04	16.7	33.3	26.7	23.3		
Audience Comfort (1 – 4)	3.45	0.57	0.0	3.3	63.3	33.3		
Central Message (1 – 4)	3.64	0.55	0.0	3.3	13.3	83.3		
Appropriate to Audience (1 – 5)	4.59	0.68	0.0	3.3	3.3	26.7	66.7	
Multisensory (1 – 3)	1.73	0.77	43.3	36.7	20.0			
Physical Engagement (1 – 4)	1.68	1.43	56.7	0.0	16.7	16.7	10.0	
Verbal Engagement (1 – 5)	3.3	1.26	0.0	46.7	20.0	6.7	26.7	
Surprise (1 – 3)	1.80	0.62	23.3	56.7	20.0			
Novelty (1 – 3)	1.95	0.72	26.7	50.0	23.3			
Consistency (1 – 3)	2.93	0.26	0.0	10.0	90.0			
Resource Quality (1 – 3)	2.02	0.80	33.3	36.7	30.0			
Multiple Viewpoints (1 – 3)	2.46	0.51	0.0	50.0	50.0			
Multiple Activities (1 – 4)	2.25	1.27	40.0	13.3	23.3	23.3		
Pace (0 – 2)	1.98	0.30	10.0	83.3	6.7			
Unexpected Negative Circumstance (0 – 1)	0.07	0.26	93.3	6.7				
Unexpected Positive Circumstance (0 – 1)	0.05	0.23	96.7	3.3				

It was expected that mean program characteristics would vary significantly by program type (traditional vs. facilitated). An independent-samples t test was performed for each characteristic (Table 11) with the higher mean for each significant difference bolded. Significant mean differences were detected for seven (25%) of the tests. Significant mean differences were found for the following seven characteristics: (1) cognitive engagement ( $t = -5.62$ ,  $p. < .001$ ), (2) relevance to audience ( $t = -2.75$ ,  $p. = .011$ ), (3) affective messaging ( $t = -3.42$ ,  $p. = .002$ ), (4) fact-based messaging ( $t = 3.20$ ,  $p. = .008$ ), (5) provocation ( $t = -5.51$ ,  $p. < .001$ ), (6) verbal engagement ( $t = -3.17$ ,  $p. = .004$ ), and (7) multiple activities ( $t = -3.07$ ,  $p. = .005$ ). For all tests except fact-based messaging, facilitated dialogue programs had significantly higher means.

For organization, no significant mean differences for the seven items were detected and thus on these program characteristics, traditional and facilitated programs were much the same. However, for the connection dimension with its seven items, five significant mean differences were found. For example, facilitated programs were rated significantly higher on cognitive engagement (mean difference = 1.54). In addition, facilitated dialogue programs were rated significantly higher in relevance to audience, affective messaging, and provocation while traditional programs were rated higher in fact-based messaging. Finally, facilitated programs were rated significantly higher in verbal engagement and multiple activities.

Table 12: Descriptive Program Characteristics by Program Type (n = 26)

Program Characteristic	Traditional Program Mean	Facilitated Program Mean	t value	p value
Organization				
• Introduction Quality (1 – 3)	2.73	2.91	-1.18	n.s.
• Appropriate Sequence (1 – 4)	3.73	3.45	1.17	n.s.
• Transitions (1 – 4)	3.73	3.73	0.03	n.s.
• Holistic Story (1 – 5)	2.40	2.27	0.31	n.s.
• Conclusion Linked to Introduction (1 – 4)	2.80	3.27	-1.35	n.s.
• Clear Theme (1 – 4)	3.87	3.73	0.87	n.s.
• Appropriate Logistics (1 – 4)	3.73	3.73	0.03	n.s.
Connection				
• Connection to Intangibles/Universals (1 – 5)	4.20	4.18	0.06	n.s.
• Cognitive Engagement (1 – 5)	2.73	<b>4.27</b>	-5.62	<b>.000</b>
• Relevance to Audience (1 – 5)	3.40	<b>4.45</b>	-2.75	<b>.011</b>
• Affective Messaging (1 – 5)	2.53	<b>3.91</b>	-3.42	<b>.002</b>
• Fact-Based Messaging (1 – 3)	<b>2.67</b>	1.55	3.20	<b>.008</b>
• Place-Based Messaging (1 – 5)	3.67	4.09	-0.99	n.s.
• Provocation (1 – 4)	1.67	<b>3.45</b>	-5.51	<b>.000</b>
Audience Comfort (1 – 4)	3.53	3.73	-0.85	n.s.
Central Message (1 – 4)	3.47	3.45	0.05	n.s.
Appropriate to Audience (1 – 5)	4.53	4.73	-0.78	n.s.
Multisensory (1 – 3)	1.80	1.55	0.81	n.s.
Physical Engagement (1 – 4)	2.13	2.27	-0.31	n.s.
Verbal Engagement (1 – 5)	2.93	<b>4.64</b>	-3.17	<b>.004</b>
Surprise (1 – 3)	1.67	1.55	0.61	n.s.
Novelty (1 – 3)	1.87	2.00	-0.44	n.s.
Consistency (1 – 3)	3.00	2.91	1.00	n.s.
Resource Quality (1 – 3)	1.93	2.27	-1.08	n.s.
Multiple Viewpoints (1 – 3)	2.00	2.75	-3.00	n.s.
Multiple Activities (1 – 4)	1.60	<b>3.00</b>	-3.07	<b>.005</b>
Pace (0 – 2)	2.00	2.00	0.00	n.s.
Unexpected Negative Circumstance (0 – 1)	0.07	0.09	-0.22	n.s.
Unexpected Positive Circumstance (0 – 1)	0.00	0.18	-1.75	n.s.

## Program Characteristics

Moreover, programs were assessed based on the coding form established by Stern, et.al. We assessed programs on 26 criteria (dimensions) using ordinal classification schemes. Total scores for the programs were summed across the 26 criteria, yielding a possible low score of 26 for a particular program and a potential high score of 104 (assuming a program received the highest possible score for each item). Total scores ranged from a low of 63 to a high of 102. There was a mean score of 79.2 (SD = 11.44) with a 95% confidence interval from 76.2 to 82.2. Individual items (all scoring at least .90) with the highest mean scores were: consistency, clear theme, introduction quality, traditions, appropriate logistics, appropriate sequence, and audience comfort. Individual items having relatively low means with a mean of less than .60 were: multisensory, physical engagement, multiple activities, surprise, and holistic story. Facilitated dialogue programs had a significantly ( $t = -2.375$ ,  $df = 24$ ,  $p = .026$ ) higher mean score ( $M = 84.8$ ) compared to the traditional programs ( $M = 75.8$ ). Significant mean differences were detected for seven program characteristics: relevance to audience, fact-based messaging, cognitive engagement, verbal engagement, affective messaging, provocation, and multiple activities. Facilitated dialogue programs had significantly higher means for six of the seven comparisons (relevance to audience, cognitive engagement, verbal engagement, affective messaging, provocation, and multiple activities), but for fact-based messaging, traditional programs had the higher mean (as should be expected).

We examined total program score by number of attendees. The number of attendees was recoded into quartiles: 1-15, 16-25, 26-41, and 42-102. Mean program

score was significantly and inversely related to number of attendees ( $F = 9.16$ ,  $df = 3$  and  $52$ ,  $p = .000$ ). The relationship was strong and negative ( $r = -.564$ ). The highest mean program score was associated with the smallest group size ( $M = 89.0$ ) while the lowest program score was associated with the largest group size ( $M = 71.2$ ). The two pairs of significantly different means were: (1) 1-15 and 26-41 and (2) 1-15 and 42-102.

Total program score was analyzed by gender of interpreter. Programs led by women scored 80.0, on average, while programs led by men averaged 77.5. The mean difference was not significant ( $t = 0.752$ ,  $df = 54$ ,  $p = .455$ ).

We then examined total program score by district in which the program was given. On average, program scores were lowest at Colter Bay ( $M = 72.4$ ) and highest at LSR Preserve ( $M = 87.00$ ). Mean differences were significant ( $F = 7.32$ ,  $df = 3$  and  $52$ ,  $p = .000$ ). The Duncan post hoc test revealed Colter Bay had a significant lower mean score than the other districts.

Next, we examined total program score by start time of the program (morning, afternoon, and evening). On average, program score was highest in the afternoons ( $M = 81.1$ ) and lowest in the evenings, but the ANOVA test was not statistically significant ( $F = 0.866$ ,  $p = .426$ ).

We wanted to know if advertised program length was associated with program score. Results indicate that while longer programs scores slightly higher than shorter program, the mean difference was not significant ( $t = -1.08$ ,  $df = 45.22$ ,  $p = .303$ ).

Finally, we hypothesized that total program score would vary significantly by year. Program score was slightly higher ( $M = 79.6$ ) in 2015 compared to 2016 ( $M = 78.8$ ), but the mean difference was not significant ( $t = 0.264$ ,  $df = 54$ ,  $p = .793$ ).

## ARC Achievement

We measured whether or not each a program included the entire ARC (4 phases). The results (Table 13) below show that 12 of the 15 (80%) traditional programs do not reach even the first step of the ARC, but one traditional program did complete the entire ARC. And, among the facilitated dialogue program, 10 of the 11 programs reached at least step 3 of the 4 while 1 program did not reach any of the ARC steps. Once again, the variables and characteristics of each program and audience strongly affect the outcome of a program. In addition, the definition of a full performance interpreter is one who knows their site content well enough and is able to assess their audience well and quickly enough to be able to turn on a dime to adjust a program accordingly to meet the needs of the situation. Adjustments in a program are expected and welcomed in order to better meet the needs and expectations of the visitors.

**Table 13: ARC Totals**

ARC Step	2015	2016
Community Building	50.0	36.7
Sharing Diversity of Experiences	38.5	26.7
Exploring Diversity of Experience	42.3	23.3
Synthesizing the Experience	34.6	20.0

**Table 14: ARC Totals by Program Type**

ARC Step	Traditional	F.D.	Traditional	F.D.
	2015	2015	2016	2016
Community Building	18.3	100.0	25.0	83.3
Sharing Diversity of Experiences	6.3	90.0	12.5	83.3
Exploring Diversity of Experience	6.3	100.0	8.3	100.0
Synthesizing the Experience	6.3	80.0	4.2	83.3

The expectation was that the traditional programs would not include the entire ARC while the facilitated dialogue programs would include the full ARC. There was a significant mean difference in number of steps in the facilitated dialogue programs ( $M = 3.4$ ) compared to the traditional programs ( $M = 0.40$ ) ( $t = -6.66$ ,  $df = 24$ ,  $p = .000$ ). In addition to testing all ARC steps, we tested each step separately to determine if significant mean differences existed between traditional programs and facilitated dialogue programs.

It was expected that significant mean differences between program types would be found for each step in the ARC model. All four inferential tests were significant (step 1:  $t = -4.811$ ,  $df = 24$ ,  $p < .0001$ ; step two:  $t = -5.786$ ,  $df = 24$ ,  $p < .0001$ ; step 3:  $t = -7.659$ ,  $df = 24$ ,  $p < .0001$ ; and step 4:  $t = -4.619$ ,  $df = 24$ ,  $p < .0001$ ). These results suggest that there are clear differences between traditional and facilitated programs with regard to their adherence or not to the ARC model. Facilitated dialogue programs were significantly more likely than traditional programs to achieve each individual step.

### **By District**

In addition, we examined ARC total by district to determine if there were significant mean differences of ARC total by district. Results indicate that significant mean differences existed ( $F = 3.922$ ,  $p = .022$ ). Jenny Lake programs, on average, had the highest ARC total ( $M = 3.80$ ) while LSR Preserve Programs had the lowest ARC total mean ( $M = 0.75$ ).

We examined selected factors to determine if they were related to ARC compliance. First, we examined the district in which each program was offered. For none of the four steps (community building, sharing diversity of experiences, exploring diversity of experiences, and synthesizing the experience) were there significant mean

differences. However, there was a significant mean difference ( $F = 3.922$ ,  $df = 3, 22$ ,  $p = .022$ ) in the ARC Total based on the district in which the program was offered. The only pairwise significant mean difference (mean difference =  $-2.709$ ,  $p = .028$ ) was between Jenny Lake ( $M = 3.80$ ) and Colter Bay ( $M = 1.09$ ). Therefore, Jenny Lake programs studied ( $n = 5$ ) had significantly higher adherence to the ARC model compared to Colter Bay. Additionally, LSR Preserve programs ( $n = 4$ ) had the lowest ARC Total ( $M = 0.75$ ).

**Table 15: ARC Total by District**

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Min	Max
					Lower Bound	Upper Bound		
					Colter Bay	11		
Jenny Lake	5	3.80	.447	.200	3.24	4.36	3	4
LSR Preserve	4	.75	1.500	.750	-1.64	3.14	0	3
Moose	6	1.50	1.975	.806	-.57	3.57	0	4
Total	26	1.65	1.853	.363	.91	2.40	0	4

Additionally, we examined ARC Total by time of day of program offering. ARC Total varied from a mean of 2.43 for morning to a mean low of 0.00 for evening programs. Significant mean differences were detected ( $F = 3.595$ ,  $df = 2$  and  $23$ ,  $p = .044$ ), but none of the pairwise mean differences were significant.

**Table 16: ARC Total by Time of Day Program**

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Min	Max
					Lower Bound	Upper Bound		
					Morning	14		
Afternoon	9	1.00	1.732	.577	-.33	2.33	0	4
Evening	3	.00	.000	.000	.00	.00	0	0
Total	26	1.65	1.853	.363	.91	2.40	0	4

Then, we investigated whether the length of the program affected the ARC Total. The ARC Total mean varied very slightly, but not significantly ( $F = 0.001, p = .975$ ).

Table 17: Length of Program affecting ARC Total :

		Correlations				
		Arc Total	Number of Attendees	Advertised Program Length	Actual Program Length	Number of Attendees Who Left Early
Arc Total	Pearson Correlation	1	-.405*	.206	.170	-.439*
	Sig. (2-tailed)		.040	.313	.405	.025
	N	26	26	26	26	26
Number of Attendees	Pearson Correlation	-.405*	1	-.025	.084	.887**
	Sig. (2-tailed)	.040		.904	.684	.000
	N	26	26	26	26	26
Advertised Program Length	Pearson Correlation	.206	-.025	1	.859**	.089
	Sig. (2-tailed)	.313	.904		.000	.642
	N	26	26	30	30	30
Actual Program Length	Pearson Correlation	.170	.084	.859**	1	.267
	Sig. (2-tailed)	.405	.684	.000		.154
	N	26	26	30	30	30
Number of Attendees Who Left Early	Pearson Correlation	-.439*	.887**	.089	.267	1
	Sig. (2-tailed)	.025	.000	.642	.154	
	N	26	26	30	30	30

\*. Correlation is significant at the 0.05 level (2-tailed).

\*\* . Correlation is significant at the 0.01 level (2-tailed).

### Relationship between ARC Total and Program Characteristics

We hypothesized a positive relationship would exist between ARC Total and program characteristics total score. This was supported with positive, very strong

correlations ( $r = .70$ ,  $p = .000$ ) for 2015 and for 2016 ( $r = .71$ ,  $p = .000$ ); programs achieving higher ARC Totals tended to have higher total scores for program characteristics.

A One-Way ANOVA test was performed to see if ARC Total was related to significant mean differences in total program scores. Significant mean differences were detected ( $F = 9.91$ ,  $df = 3$  and  $22$ ,  $p < .0005$ ). Programs having a total ARC score of 0 had a significantly lower mean ( $M = 73.8$ ) compared to those achieving an ARC score of 4 ( $M = 91.1$ ).

ARC Total was regressed on total program score. Almost 50% ( $R^2 = .49$ ) of the variability in total program score was explained through its association with ARC total. In addition, the regression equation was: Total Program Score =  $73.135 + 3.918$  ARC Total. This means that each for each step increase in ARC, there is a predicted increase of almost 4 points in the total program score.

## **Conclusions: Part I**

During Summers 2015 and 2016, 56 traditional and facilitated dialogue interpretative programs conducted in Grand Teton National Park were studied using unobtrusive observation of park interpreters and their programs. All programs were recorded and all programs were studied and evaluated using a data collection form. Attending these programs were 1,147 individuals with number of attendees ranging from a low of four to a high of 128. Programs had a mean of 29 attendees with a median of 19. Attendees were more likely to be female (57%), adult (67%), and white (89%). No attendees left in 37.5% ( $n = 21$ ) of studied programs. One or more attendees left their

program before program's end of most programs. Approximately 10% of initial attendees left the program before program's end ( $M = 3$  per program). Eleven programs were conducted in the Colter Bay area while there were five at Jenny Lake, 4 at LSR Preserve, and 6 at Moose. Fifteen of the 26 programs were identified as traditional programs and 11 were identified as facilitated dialogue. Over half ( $n = 14$ ) of the 26 programs studied were in the morning. Type of program varied somewhat (but not significantly) by district with 80% of programs at Jenny Lake identified as facilitated dialogue and only a third of programs at Colter Bay also identified as facilitated dialogue. Advertised length of programs varied from 20 minutes to 150 minutes ( $M = 47.7$  minutes). Some 54% of programs were advertised as being 30 minutes or less.

Comparing traditional programs to facilitated dialogue questions, we found three significant differences: number of attendees, number of males and number of whites. Traditional programs, on average, had more attendees, more males, and more whites compared to facilitated dialogue programs. Twenty-seven characteristics were studied using an ordinal classification system for each program (see Appendix A). Generally, programs appeared to be delivered professionally by competent interpreters.

In terms of ARC adherence, the eleven facilitated dialogue programs averaged 3.4 of the 4 steps. Three programs made it to step 3 while 7 of the programs made it all the way through step 4. There were significant mean differences between facilitated dialogue and traditional programs, suggesting that interpreters were approaching and delivering programs appropriately based on program type.

Programs were assessed based on the coding form established by Stern, et.al. We assessed programs on 26 criteria (dimensions) using ordinal classification schemes.

Furthermore, we hypothesized that the two program types would be differentiated on the basis of program characteristics. Total scores for the programs were summed across the 26 criteria, yielding a possible low score of 26 for a particular program and a potential high score of 104 (assuming a program received the highest possible score for each item). Total scores ranged from a low of 65 to a high of 99. There was a mean score of 79.6 (SD = 10.42) with a 95% confidence interval from 75.4 to 83.8. Individual items (all scoring at least .90) with the highest mean scores were: consistency, clear theme, introduction quality, traditions, appropriate logistics, appropriate sequence, and audience comfort. Individual items having relatively low means with a mean of less than .60 were: multisensory, physical engagement, multiple activities, surprise, and holistic story. Facilitated dialogue programs had a significantly ( $t = -2.375$ ,  $df = 24$ ,  $p = .026$ ) higher mean score ( $M = 84.8$ ) compared to the traditional programs ( $M = 75.8$ ).

Significance tests were performed for all 26 characteristics to determine if significant mean differences existed between traditional and facilitated dialogue programs. No significant differences were detected for 19 program characteristics, indicating consistency between program types for about 73% of the tested program characteristics. However, significant mean differences were detected for seven program characteristics: relevance to audience, fact-based messaging, cognitive engagement, verbal engagement, affective messaging, provocation, and multiple activities. Facilitated dialogue programs had significantly higher means for six of the seven comparisons (relevance to audience, cognitive engagement, verbal engagement, affective messaging, provocation, and multiple activities), but for fact-based messaging, traditional programs had the higher mean (as should be expected).

## **Results: Part II – Participant Survey**

During Summer 2016, Interpreters giving interpretive programs in the park were asked to solicit for survey participation at the end or beginning of their programs using a pre-written script. Email addresses were obtained at these times from adults willing to participate. In late summer, email addresses were entered into Qualtrics so that willing participants could receive the hyperlink to the OMB approved online survey. Multiple reminders were emailed to those who signed up for the survey. The survey opened on September 15<sup>th</sup> and closed November 6<sup>th</sup>; the survey was open for 52 total days. After the end of the survey deadline, the data file was downloaded from Qualtrics into SPSS for analysis. Data from a total of 168 useful responses was analyzed using SPSS.

### **Demographic Characteristics**

We received a total of 168 responses from visitors solicited for the survey. Almost 63% (n =98) of respondents were women while 37% (n =58) were men. Respondents' mean age was 51.7 (SD 13.4). The youngest was 17 while the oldest was 77. Seventy-five percent were age 63 or younger while 25% were older than 64. Over 97% self-identified as non-Hispanic while 97% also self-identified as white. The modal and median educational level was a graduate or professional degree. The typical respondent (78.8%) was in the Park with family members. And, 94.1% of respondents were residing in the United States while 5.9% were residing outside the United States. Of these 9 respondents, three were from The Netherlands, two from Canada, and one each from Germany, United Kingdom, Israel, and South Africa. Lastly, some 41.3% of respondents participated in the Junior Ranger Program and earned the Badge.

## **GRTE Visits**

For just over half (52.4%), reported that this visit was their first visit to Grand Teton National Park. For all respondents mean number of visits was just over 1 ( $M = 1.42$ ,  $SD = 3.25$ ), but among only those who had visited the Park prior to this summer the mean was 3.14 ( $SD = 4.25$ ). Only about 10% of respondents had visited the Park in three or more times and for one respondent in seven this was only their second park visit.

## **Opinions Concerning the Park**

Respondents were asked how they felt or what they believed about 13 statements which measured place attachment and concerned Grand Teton National Park. Answers ranged from strongly disagree (1) to strongly agree (5). Three statements (Table 18) all had mean ratings of agree (4.0) or higher: "Grand Teton National Park means a lot to me," "I am very attached to Grand Teton National Park," and "Grand Teton National Park is a special place for my family." Two other statements ("Many important family memories are tied to Grand Teton National Park" and "I identify strongly with Grand Teton National Park") were slightly below 4.0.

Table 18: Descriptive Statistics of Opinions Toward the Park

	Mean	Std. Deviation
Grand Teton National Park means a lot to me.	4.33	.930
I am very attached to Grand Teton National Park.	4.13	.822
Grand Teton National Park is a special place for my family.	4.09	.737
Many important family memories are tied to Grand Teton National Park.	3.97	.974
I identify strongly with Grand Teton National Park.	3.97	.877
I feel a sense of pride in my heritage when I am at Grand Teton National Park.	3.81	.802
I feel that I can really be myself at Grand Teton National Park.	3.80	.853
I feel Grand Teton National Park is part of me.	3.67	.873
Visiting Grand Teton National Park says a lot about who I am.	3.61	.823
I get more satisfaction out of visiting Grand Teton National Park than from visiting any other similar parks.	3.42	.871
No other place can compare to Grand Teton National Park.	3.38	.837
Few people know Grand Teton National Park like I do.	2.65	1.026
I feel no connection to Grand Teton National Park.	1.42	.761

These statements were factor analyzed using principal components analysis with varimax rotation and with factor loadings of .60 or greater retained. Three components emerged, which explained 63.03% of variance. The first factor, Identity, included three variables: "Visiting Grand Teton National Park says a lot about who I am," "Grand Teton National Park is a special place for my family," and "Many important family memories are tied to Grand Teton National Park." The second factor, Park Attachment, included four variables: "Grand Teton National Park means a lot to me," "I am very attached to Grand Teton National Park," "I identify strongly with Grand Teton National Park," and "I feel no connection to Grand Teton National Park." And, the third component, Park Uniqueness, consisted of two variables: "I get more satisfaction out of visiting Grand

Teton National Park than from visiting any other similar parks” and “No other place compare to Grand Teton National Park.”

### Programs Attended

We asked respondents on a scale of very unimportant (1) to very important (5) to indicate their level of importance they place on selected factors when choosing interpretive programs to attend. Mean importance ratings varied from a high of 4.21 for to a low of 3.50. Four factors (time, location, day, and content) all had mean score above 4.0 with time receiving the highest mean. Program title (M = 3.52) and program format (M = 3.50) were considerably lower than the first four factors in importance.

**Table 19: Importance of Selected Factors in Choosing Programs**

	N	Mean	Std. Deviation
Program Time	150	4.21	.922
Program Location	152	4.18	.872
Program Day	148	4.09	.914
Program Content	153	4.07	.717
Program Title	152	3.52	.996
Program Format	147	3.50	.855

We factor analyzed (varimax rotation) these six program selection criteria to determine which factors hung together in their thinking. Two principal components emerged, which combined to explain 63.2% of variance. The first factor, practicality, included day of program, program time, and program location. The second factor included program format and content items. Interestingly, program title did not load into either of the two principal components.

Some 71.0% of respondents said they attended at least one traditional ranger-led program during their visit. Number of traditional programs attended ranged from a low of 1 (60.5%) to a high of 20 (0.9%). Respondents averaged attending about two programs (SD = 2.6) during their GRTE visit. We then asked if they attended any of the “new type” (facilitated dialogue) programs while in the Park. Exactly 50% said they attended at least one. Most (77.2%) attending at least one said they attended only one. Another 13% attended two such programs. Mean number of conversational-type programs attended was 1.4 (SD = 0.88). Only 7.8% said they knew the program was a new type of program before they attended.

We then compared the number of traditional programs attended to the number of facilitated dialogue programs attended. Twelve percent of respondents attended more facilitated dialogue programs than traditional programs, 47% attended the same number of each, and 41% attended more traditional programs.

Number of programs attended for the two types were summed to produce a total number of programs attended during the visit. On average, these respondents reported attended almost four programs ( $M = 3.77$ ,  $SD = 3.18$ ); the 95% confidence interval was 2.93 – 4.62 programs. We then analyzed total number of programs attended by selected predictors. While men averaged slightly more programs compared to women, the mean difference was not statistically significant. On average, those in their forties attended the most programs while senior adults attended the fewest; the mean difference was not significant. Those with a graduate or professional degree averaged the most programs ( $M = 4.21$ ) while those without a bachelor's degree attended the fewest ( $M = 2.5$ ); once again the mean difference was not statistically significant.

Repeat visitors attended slightly more programs than first-time visitors, but no statistically significant difference was detected.

We then asked those attending programs which type they preferred. Over 56% said they preferred the traditional program while 44% said they preferred the conversational program type. We then looked for relationships between selected predictor variables and program type preference. No significant relationships were found for the following predictors: (1) gender with males slightly more likely to prefer conversational programs; (2) age with those in their 40s slightly more likely to prefer conversational programs; (3) race with whites slightly more likely to prefer conversational programs; (4) education with college graduates slightly more likely to prefer conversational programs; (5) country of origin with those from the United States slightly more likely to prefer conversational programs. (6) previous visit to GRTE with repeat visitors more likely to prefer conversational programs.

As a follow-up question, we asked respondents to indicate the amount of audience conversation they prefer during programs. Answers ranged from very little to a great deal (Table 20). Only one-in-seven said they prefer very little audience conversation and about the same percent said, "quite a bit" or "a great deal." Median response was "some."

Table 20: Amount of Audience Conversation Preferred

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Very Little	28	16.7	17.4	17.4
	Some	104	61.9	64.6	82.0
	Quite a Bit	25	14.9	15.5	97.5
	A Great Deal	4	2.4	2.5	100.0
	Total	161	95.8	100.0	
Missing	No Answer	7	4.2		
Total		168	100.0		

The Junior Ranger Program is a mainstay in national parks and GRTE is no exception. We asked respondents if they participated in the program during their park visit (Table 21).

Table 21: Participate in the Junior Ranger Program

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes, and earned the Junior Ranger Badge	69	41.1	41.3	41.3
	Yes, participated but did not complete	5	3.0	3.0	44.3
	No, we did not have any children in the group	75	44.6	44.9	89.2
	No, we had children in the group, but did not participate	18	10.7	10.8	100.0
	Total	167	99.4	100.0	
Missing	No Answer	1	.6		
Total		168	100.0		

Some 41.3% said they did participate and someone in their group earned a badge while another 3% said they participated but a badge was not received. And,

importantly, some 10.8% said they had children in their group, but did not participate in the program.

Which programs did they say they attend? Results in Table 22 are in descending order with the proportion of respondents indicating they attended a program. Thus, .24, or 24%, of respondents said they attended the Taggart Lake Hike from Moose. Of special note, is that the top four programs all involved some walk or hike associated with the program.

**Table 22: Program Attendance in Descending Order**

	Mean	Std. Deviation
Moose: Taggart Lake Hike	.24	.431
Jenny Lake: Inspiration Point	.17	.374
LSR Preserve: Explore the Preserve Hike	.14	.345
Moose: A Walk into the Past	.13	.338
Colter Bay: Campfire Program	.10	.302
Moose: Your Park, Your Legacy	.10	.302
Moose: Twilight Talk	.07	.258
Moose: Bear Safety	.07	.248
Moose: Nature in a Nutshell	.06	.237
Colter Bay: Swan Lake Hike	.05	.214
Colter Bay: Grand Teton Kids	.05	.214
Jenny Lake: Bear Safety	.04	.200
Moose: Map Chat	.04	.200
Colter Bay: Early Evening	.04	.186
Colter Bay: Celebrating National Parks	.04	.186
Jenny Lake: Discover Grand Teton	.04	.186
LSR Preserve: Start with a Ranger	.04	.186
LSR Preserve: Critter Chat	.03	.170
Colter Bay: Teton Highlights	.02	.153
LSR Preserve: Bear Safety	.02	.153
Colter Bay: Indian Arts and Culture	.02	.153
Colter Bay: Tipi Demo	.02	.153
Colter Bay: Morning on the Back Deck	.02	.133
Colter Bay: Bear Safety	.02	.133

LSR Preserve: Nature Explorer's Backpack	.01	.109
Colter Bay: Coffee with a Ranger	.01	.077
Colter Bay: Family Night Video	.01	.077

### Impact of Programs Attended

We asked respondents to indicate the degree to which the program or programs attended influence aspects of their thinking, beliefs, opinions, or behavior for thirteen items. Answers ranged from a low of “not at all” (1), to “a little” (2), “somewhat” (3), “moderate amount” (4) and to a high of “a great deal” (5). Results are shown in Table 23 below in mean descending order. Eleven of the 13 means were above 4 (moderate amount). Only two items (“Changed the way I will behave after I leave this park” and “Changed the way I will behave while in this park”) had means below 3 (somewhat). These may have been relatively low because behavioral change was not a primary focus of park programs. “Made my visit to this park more enjoyable” (M = 4.43) and “increased my knowledge about the program’s topic” (M = 4.41) had the highest means.

**Table 23: Factors Influencing Attitude or Behavioral Outcome in Descending Order ALL Programs**

	Mean	Std. Deviation
Made my visit to this park more enjoyable	4.43	.840
Increased my knowledge about the program's topic	4.41	.868
Enhanced my appreciation for this park	4.39	.864
Enhanced my appreciation for the National Park Service	4.38	.919
Made my visit to this park more meaningful	4.32	.915
Made me care more about protecting places like this	4.09	1.227
Made me want to tell others about what I learned	3.95	1.133
Made me care more about this park's resources	3.83	1.230
Made me more likely to avoid harming park resources	3.79	1.416
Made me think deeply	3.72	1.090
Made me reflect on my own life	3.23	1.194
Changed the way I will behave after I leave this park	2.94	1.366
Changed the way I will behave while in this park	2.86	1.409

**Table 24: Factors Influencing Attitude or Behavioral Outcome in Descending Order Participants Attending at Least One Facilitated Dialogue Program**

	Mean	Std. Deviation
Made my visit to this park more enjoyable	4.61	.655
Increased my knowledge about the program's topic	4.57	.772
Enhanced my appreciation for this park	4.51	.788
Made my visit to this park more meaningful	4.45	.839
Enhanced my appreciation for the National Park Service	4.43	.895
Made me care more about protecting places like this	4.09	1.277
Made me want to tell others about what I learned	4.07	1.064
Made me think deeply	3.89	1.014
Made me care more about this park's resources	3.86	1.243
Made me more likely to avoid harming park resources	3.84	1.461
Made me reflect on my own life	3.42	1.218
Changed the way I will behave after I leave this park	3.22	1.446
Changed the way I will behave while in this park	3.08	1.487

In addition, we performed a principal components analysis with varimax rotation of these 13 items to determine how many theoretical dimensions emerged from the data patterns. Combined these two components explained 69.3% of variance. Loading into the first factor (Appreciation and Meaningfulness) with loadings greater than .70 were six items: made my visit to the park more enjoyable, increased my knowledge about the program topic, enhanced my appreciation for this park, made my visit to this park more meaningful, and enhanced my appreciation for the National Park Service. Loading into the second factor (Care and Change) were five factors with factor loadings of greater than .60: changed the way I will behave after I leave the park, changed the way I will behave while in the park, made me more likely to avoid harming park resources, made me care more about this park's resources, and made me care more about protecting places like this. Next, we examined mean rating differences for these items based on respondent's preference for traditional programs or facilitated dialogue programs.

**Table 25: Comparison of Traditional Programs with Facilitated Dialogue Programs on Influence of Attitude and Behavior**

Item	Traditional Program Mean	Facilitated Dialogue Program Mean	Mean Difference	Significance
Made me think deeply	3.49	4.05	-0.554	.002
Made me reflect on my own life	2.94	3.63	-0.687	.001
Enhanced my appreciation for this park	4.30	4.59	-0.297	.025
Enhanced my appreciation for the National Park Service	4.28	4.58	-0.294	.041
Made me more likely to avoid harming park resources	3.46	4.29	-0.829	<.0005
Increased my knowledge about the program's topic	4.32	4.57	-.250	n.s.
Made my visit to this park more enjoyable	4.35	4.65	-0.299	.016
Made my visit to this park more meaningful	4.14	4.59	-0.451	.002
Changed the way I will behave while I'm in this park	2.58	3.25	-0.666	.005
Changed the way I will behave after I leave this park	2.60	3.40	-0.803	<.0005
Made me want to tell others about what I learned (	3.79	4.27	-0.487	.007
Made me care more about this park's resources	3.69	4.06	-0.375	n.s.
Made me care more about protecting places like this	3.96	4.38	-0.443	.029

Of the 13 comparisons, eleven mean differences were statistically significant (Table 24). For each test, facilitated dialogue programs received the higher mean. For example, facilitated dialogue program preference had a significantly higher mean for “made me think deeply” compared to traditional program preference.

## **Conclusions: Part II**

Grand Teton National Park is a popular vacation destination for individuals and families. Furthermore, it is clear that park interpreters can and should play a vital role in helping visitors to enjoy their park visit, to learn about the park natural world and the broader natural world. Based on this portion of the research it is clear that there is a place for both traditional and facilitated dialogue/audience centered programming and, as a matter of fact, this may help with drawing in a more diverse visitor representation to hear the messages of the park. Though, in this study, we were unable to capture behavior change, we feel a longer term study may be able to measure that influence on visitors.

In addition, we want to make sure you know of the numerous positive comments about the programs. One respondent sent a separate email because there was not a question asking about the quality of the programs and they had wanted to make sure we knew how much they had enjoyed and appreciated the talent of your interpreters.

As far as the Arc of Dialogue is considered in the study, there was clearly growth with the practice over two summers. The second summer we observed more relaxation of the interpreter in using the technique and more use of some of the phases integrated into traditional programs as well. The Arc is certainly only one design of facilitation so

variations of the concept of audience-centered viewed in the second summer were a welcomed observed growth. Since audience-centered programming is best when the audience and purpose are the first components of consideration and because GRTE prides itself on being on the cutting edge of the field, it not surprising to see these changes. In addition, we observed several of the traditional programs incorporating audience-centered and facilitated dialogue techniques. We expect that ultimately there will be three categories with somewhat blended lines of demarcation: traditional, hybrid, and facilitated dialogue. We encourage GRTE to continue on this path.

### **Recommendations and Considerations**

Our recommendations are:

- GRTE staff should continue both traditional and facilitated dialogue programs. It's clear these two program types both have value and impact.
- GRTE staff should continue to conduct an attendee response study. This study should not only measure participant perceptions of and opinions toward both program types, but analysis should be conducted to better determine program type strengths and weaknesses. In addition, if GRTE and NPS truly want to impact behavior and attitude change, a more longitudinal study is needed to confirm or dispel changes.
- GRTE should consider the fact that the two program types may be drawing different population types. Whether or not GRTE desires different participant types is a question that should be addressed.

- GRTE staff should consider the fact that the four districts did not offer facilitated dialogue programs in the same relative proportions. Jenny Lake had the greatest proportion of facilitated dialogue programs while Colter Bay exhibited the smallest proportion of facilitated dialogue programs. Should there be more uniformity? This is merely a question to be considered based on the character of the district and its visitors.
- Attendee persistence to program's end may be a problem to be addressed. Almost 10% of original program attendees left before the end of programs. The percent was slightly, but not significantly, higher for traditional versus facilitated dialogue programs. This percentage also varied somewhat by advertised length of program with proportionally more attendees leaving for longer programs.
- GRTE should evaluate marketing of facilitated dialogue programs, which should include names, topics, locations, and topics. Three of the facilitated dialogue programs did not "make." This represents about 20% of facilitated dialogue programs studied. It may be necessary to reevaluate which programs, taking into consideration all variables such as time, content, etc., work better as facilitated, traditional, or hybrid.
- Also in the marketing strain, GRTE may wish to reevaluate the titles of the programs to consider giving them different titles that may be more descriptive and appealing to potential audiences.
- Develop more family activities, which was a theme emerging for the survey. Families want to visit the park and develop memories. These could include more ranger-led hikes, fishing, backpacking, and other activities.

- Consider altering park programs based on these results. The keys to park selections were, in large part, practical reasons such as time of day. Some park programs are poorly attended. Consider eliminating those programs or reducing the number of offerings so staff can be deployed for other programs or in different ways. Consider having more public meetings with park visitors for senior staff to discuss park issues, challenges, and solutions.
- Continue to provide opportunities for student participation in research and interpretation. Students participating in this project have truly had transformative experiences which will make them better researchers and professionals in the field.
- The methodology used for this multi-faceted study is an appropriate template to share and use at other parks. Any one of the pieces may be broken out and utilized, however, the observations based on Stern's work, the videography, the visitor survey, and MBTI assessment of the interpreters all added to help build a more complete picture of facilitate dialogue integration and how it may continue to grow as process to not only benefit the park, but also the visitor and their home communities.
- Given the potential of the experience of facilitated dialogue to change its participants, we suggest that GRTE consider continued research in this area on the changes it may instigate in the interpreters themselves. Most of the interpreters are seasonal so it would be interesting to find out how these skills are being used outside of the park. Since one of the reasons behind incorporating facilitated dialogue was in hope of providing the participant with

experience in skills that would transcend the park visit and carry over into community action, the interpreters would be a natural place from which to start capturing that proof.

- Conduct more visitor research. Parks are people and society and both change over time.

## References

- Alderson, W.T. and S.P. Low. (1985). *Interpretation of Historic Sites* (2<sup>nd</sup> edition). Nashville, TN: American Association for State and Local History.
- Beck, Larry and T. Cable. (1988). *Interpretation for the 21<sup>st</sup> Century: Fifteen Guiding Principles for Interpreting Nature and Culture*. Sagamore Publishing.
- Knudson, Douglas M., Ted T. Cable, and Larry Beck. (1995). *Interpretation of Cultural and Natural Resources*. State College, PA: Venture Publishing, Inc.
- Mills, Enos. A. (1920). *The Adventures of a Nature Guide*. Garden City, NY: Doubleday, Page and Co.
- Shankland, R. (1970). *Steve Mather of the National Parks*. New York: Alfred A. Knopf.
- Stern, M J., Robert Powell, Emily Martin, and Kevin McLean. (2012). *Identifying Best Practices for Live Interpretative Programs in the United States Park Service*. Research Report for the U.S. National Park Service.
- Tardona, Jayne H. and Daniel R. Tardona. (1997). Natural Resource Interpretation and Conversation Education in a Global Society. *The George Wright FORUM*. Vol. 14, No. 4, pp. 44-50.
- Tilden, Freeman. (1957). *Interpreting our Heritage: Principles and Practices for Visitor Services in Parks, Museums, and Historic Places*. Chapel Hill, NC: University of North Carolina Press.
- Tilden, Freeman. (1967). *Interpreting our Heritage: Principles and Practices for Visitor Services in Parks, Museums, and Historic Places*. (revised edition). Chapel Hill, NC: University of North Carolina Press.
- United States National Park Service. (2014). *The Arc of Dialogue*. Retrieved from <http://www.interpnet.com>.
- United States National Park Service. (2014). *What is Facilitated Dialogue*. Retrieved from <http://www.interpret.net>.

**Appendix A**  
Program Scoring Form

GRAND TETON NATIONAL PARK  
INTERPRETATION PROGRAM STUDY

Program characteristics observed in the study, their definitions, & operationalization

Program characteristic	Definition	Scoring	Score
<b>Introduction quality</b>	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>	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>	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>	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>	Degree to which the program followed a logical sequence.	4 = Enhanced messaging 3 = Appropriate 2 = Choppy 1 = Detracted from messaging	
<b>Transitions</b>	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	
<b>Intangibles and universals</b>	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	5 = Extensively developed; powerful concepts 4 = Well developed 3 = Present but weak 2 = Difficult to detect or slightly used 1 = Clearly not present	

<b>Multisensory</b>	Degree to which the program intentionally and actively engaged more than just basic sight and sound.	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)	
<b>Physical engagement</b>	Degree to which the program physically engaged audience members in a participatory experience; i.e., through touching or interacting with resource.	4 = Central programming element 3 = Occurred multiple times 2 = Minimal effort to engage 1 = No efforts 0 = Not appropriate	
<b>Verbal engagement</b>	Degree to which the program verbally engaged audience members in a participatory experience; i.e., dialogue (a two-way discussion).	5 = Central programming element 4 = Most of the Time 3 = Modestly engaged 2 = Minimal effort to engage 1 = No efforts	
<b>Cognitive engagement</b>	Degree to which the program cognitively engaged audience members in a participatory experience beyond simply listening; i.e. calls to imagine something, reflect, etc.	5 = Central programming element 4 = Occurred multiple times 3 = Sometimes occurred 2 = Minimal effort to engage 1 = No efforts	
<b>Multiple activities</b>	Degree to which the program consisted of a variety of activities and opportunities for direct audience involvement (not including dialogue).	4 = 2+ primary activities included 3 = 2+ secondary activities included 2 = One secondary activity included 1 = One activity only	
<b>Props</b>	A visual aide beyond a screen-based slideshow.	1 = Prop(s) used 0 = Not used	
<b>Relevance to audience</b>	Degree to which the program explicitly communicated the relevance of the subject to the lives of the audience.	5 = Major focus of messaging 4 = Well developed efforts 3 = Moderate efforts 2 = Minimal efforts 1 = No efforts	
<b>Affective messaging</b>	Degree to which the program communicated emotion (in terms of quantity, not quality).	5 = Central programming element 4 = Frequent and repeated messages 3 = Occasional messages 2 = Minimal effort to include messages 1 = Messages absent	

<b>Fact-based messaging</b>	Degree to which the program communicated factual information.	3 = Messaging was solely fact-based 2 = Most of the time fact-based 1 = Some of the time fact-based	
<b>Surprise</b>	Degree to which the program used the element of surprise in communication. This could include "aha" moments or unexpected or contrasting messages.	3 = Major element 2 = Minor element 1 = Not used	
<b>Novelty</b>	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.	3 = Major element 2 = Minor element 1 = Not used	
<b>Provocation</b>	Degree to which the program explicitly provoked participants to personally reflect on content and its deeper meanings.	4 = Powerful and explicit inclusion 3 = Occasional inclusion 2 = Isolated or vague inclusion 1 = No attempt made	
<b>Multiple viewpoints</b>	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?)	3 = Multiple viewpoints developed; none given clear priority 2 = Primarily one viewpoint, with some focus on others 1 = No effort NA = not applicable	
<b>Holistic storytelling</b>	Degree to which the program aimed to present a holistic story (with characters and a plot) as opposed to disconnected pieces of information.	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	
<b>Place-based messaging</b>	Degree to which the program emphasized the connection between the visitor and the site/resource.	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	

<b>Introduction and conclusion linkage</b>	Degree to which program connected conclusion back to the introduction in an organized or cohesive way (i.e., program "came full circle")	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	
<b>Clear theme</b>	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.	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	
<b>Central message</b>	Degree to which program's message(s) was clearly communicated; i.e., the "so what?" element of the program.	4 = Clearly communicated and well developed 3 = Easy to detect, but not well developed 2 = Difficult to detect, ambiguous 1 = Unclear/not present	
<b>Consistency</b>	Degree to which the program's tone and quality were consistent throughout the program	3 = Consistent 2 = Some shift in either tone or quality during the program 1 = Shift in both tone and quality	
<b>Pace</b>	Degree to which the pace of the program allowed for clarity and did not detract from the program.	3 = Too fast 2 = Just fine 1 = Too slow	
<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>	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>	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>	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>	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>	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>	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	

**Appendix B**  
Data Collection Form for Program Demographics

Program Demographics	
Name of Program	
Type of Program	<input type="checkbox"/> Facilitated <input type="checkbox"/> Hybrid <input type="checkbox"/> Traditional
Arc of Dialogue (Select all appropriate boxes)	<input type="checkbox"/> Community Building <input type="checkbox"/> Sharing Diversity of Experiences <input type="checkbox"/> Exploring Diversity of Experience <input type="checkbox"/> Synthesizing the Experience
Date of Program	
Day of Program	
Time of Program	
Location of Program	
Name of NPS Interpreter	
Advertised Length of Program	
Actual Length of Program	
Total Number of Attendees	
Number of Attendees who Left Before End	
Number of Males	
Number of Females	
Number of Adults	
Number of Children	
Number of Whites	
Number of Asians	
Number of Blacks	
Weather	

Notes

