Is It the Program or the Interpreter? Modeling the Influence of Program Characteristics and Interpreter Attributes on Visitor Outcomes

Robert B. Powell
Department of Parks, Recreation and Tourism Management and School of Agricultural and Forest Environmental Sciences, Clemson University

Marc J. Stern
Department of Forest Resources and Environmental Conservation, Virginia Tech

Abstract
This study modeled the relative influence of program characteristics and interpreter attributes on three visitor outcomes (satisfaction, visitor experience and appreciation, and behavioral intentions) using Structural Equation Modeling (SEM). The three resulting models accounted for between 10% and 27% (R^2) of the variance in the outcomes. The models suggest that both program and interpreter characteristics, as well as other unaccounted for factors, influence these outcomes. We discuss the implications of the findings for researchers and practitioners, calling for greater attention to both interpreter attributes and context.

Keywords
interpretation, communications, structural equation modeling, communication theory

Introduction
Much has been written regarding the interpretive techniques that should be employed to enhance visitor outcomes (e.g., Ham, 1992, 2013; Moscardo, 1999; Knudson, Cable, & Beck, 2003; Brochu & Merriman, 2002; Lewis, 2005). These interpretive techniques, which we refer to as program characteristics, are the focus of training efforts offered by organizations such as the National Park Service (see NPS, 2003a,b,c,d,e,f) and the National Association for Interpretation (NAI) as well as college courses offered around the world. These characteristics are believed to improve the quality of interpretive communications and to contribute to reaching desired outcomes, such as inspiring audiences to form intellectual and emotional connections with interpreted resources, influencing attitudes, and in some cases motivating behaviors. Researchers and field interpreters also recognize that there are

other factors, such as the attributes of the interpreter (confidence, charisma, enthusiasm, passion, apparent knowledge, etc.) that may influence the effectiveness of an interpretive program (e.g., Ham & Weiler, 2002a). However, these attributes are often overlooked in research and interpretive training. Stern and Powell (article 1, this issue) and Powell and Stern (article 3, this issue) investigated live interpretation programs provided by the U.S. National Park Service (NPS) to examine the relationship between 56 different interpretive practices and interpreter attributes and visitor satisfaction, enjoyment and appreciation for resources, and behavioral intentions. The results suggest that only certain program characteristics and interpreter attributes were significantly related with these outcomes. This study seeks to extend these findings by modeling the relative influence of these program characteristics and interpreter attributes on visitor outcomes using structural equation modeling (SEM).

The interpretive techniques promoted by professional associations and organizations have evolved over many decades and are based on experience, expert consensus, theory, and peer-reviewed research (Skibins et al., 2012). However, the empirical support for many of these best practices is largely anecdotal, because few studies to-date have attempted to isolate the influence of particular practices on outcomes through the use of experimental (or quasi-experimental) designs or comparative approaches (Skibins et al., 2012). The isolation of the influence of particular practices is challenging even with these designs, as program outcomes inevitably emerge from a dynamic interaction between the interpreter, the audience, the content, the setting/context, and the delivery (Powell et al., 2009, 2012; Archer & Wearing, 2003; Wearing & Wearing, 2001). Accounting for all factors seems a near impossibility. In this paper, we explore the relative influence of two of these elements, interpreter characteristics and program characteristics, on visitor outcomes. We define interpreter characteristics as those that may be entirely unique to the individual interpreter in any given context. These elements might include their mood, personality, or particular style of presentation. While program characteristics may also be highly dependent upon the interpreter, they could also be incorporated by design into a pre-packaged program, such as the sequence, content, theme, or logistics of the program.

We use structural equation modeling (SEM) for two reasons. First, the models give us a sense of the relative strength of influence of interpreter and program characteristics on visitor outcomes. The models can reveal the percentage of the observed variance in each outcome that can be explained by the predictors (Byrne, 2006). Second, the models allow for an examination of the interactions between interpreter characteristics and program characteristics. SEM also reveals the most parsimonious causal models for each outcome. As such, only the most predictive combination of variables remains in the final models. Examining which variables are present in the final models and their inter-relationships allows for consideration of the relative influence of program design vs. interpreter attributes. For example, if only interpreter characteristics are present in the final models, we would consider them dominant drivers of visitor outcomes. If both interpreter and program characteristics are present, it would support a view that outcomes are produced more by the interactions between interpreter and program design rather than by one or the other.

Methods

Selection of sites
We observed 376 diverse interpretive programs provided by 24 NPS park units across six regions of the NPS that generally reflected the current makeup of the NPS system (see Stern and Powell, this issue). The criteria for selecting NPS units included:

- Annual visitation greater than 35,000
- Geographic distribution across the county
- Variable distances from urban centers (urban, urban proximate, remote)
- Resource-base (cultural, natural, mixed)
- The ability to observe multiple programs in a short period of time
- Willingness to participate

The 24 selected units varied widely in terms of visitation, resource base, and locations, providing a reasonable sample from which to make generalizations regarding interpretation provided across the NPS system.

Sampling and data collection
Four researchers collected field data. Prior to each program one researcher conducted a short interview with the interpreter to collect demographic and background information regarding the program. During the program, this same researcher monitored 56 different program and interpreter characteristics and recorded these details on standardized observation sheets. After the program, we surveyed attendees that were age 15 or older using a standardized questionnaire. For programs with fewer than 50 participants, we attempted a census of all eligible attendees. In programs with more than 50 attendees, we systematically sampled attendees. From the 376 programs, we collected 3,603 surveys from visitors (for more detail, see Stern & Powell, this issue).

Data cleaning
Post-program surveys and program audits were coded and entered into Microsoft Access Database and Microsoft Excel to facilitate data entry. Data were then transferred to SPSS and EQS v6.1 software (Bentler, 2005) for screening and analyses. The visitor survey data were first screened for cases missing more than 50% of the items per factor (Tabachnick & Fidell, 2007). A total of 118 respondents were removed as a result. Data were then screened for univariate and multivariate outliers on outcome variables following Tabachnick and Fidell (2007) using Mahalanobis Distance (MAH) and studentized deleted residuals (SDRESID). A total of 58 cases were removed for exceeding +/- 3 standard deviations, or the criterion Mahalanobis Distance value (Fox 1991). This reduced our sample to 3,427 individual surveys from 376 interpretive programs.

Next we reviewed the number of valid respondents per individual interpretive program. Prior theory and research suggest that programs with a low number of attendees may be inherently different than programs servicing a larger number of attendees (Forist, 2003; McManus, 1987, 1988; Moscardo, 1999; Stern & Powell, this issue). We observed 272 programs with five or more attendees (see Stern & Powell,

Table 1. Outcome (dependent) variables with descriptive statistics.

Outcomes	N	Mean	St. Dev.
Satisfaction: 0 to 10 scale	272	8.94	0.64
Visitor experience and appreciation **(α=.89): 1 to 5 scale**	272	4.41	0.32
• Made my visit to this park more enjoyable		4.55	0.30
• Made my visit to this park more meaningful		4.49	0.32
• Enhanced my appreciation for this park		4.36	0.37
• Increased my knowledge about the program's topic		4.45	0.34
• Enhanced my appreciation for the National Park Service		4.27	0.36
Behavioral intentions **(α=.94): 1 to 5 scale**	272	2.92	0.64
• Changed the way I will behave while I'm in this park		2.92	0.67
• Changed the way I will behave after I leave this park		2.92	0.61

article 1, this issue for more extensive description). We chose this sample for the analyses conducted herein because it is most representative of programs in general and it provides a sample large enough to conduct structural equation modeling (Byrne, 2006). Because the program was our unit of analysis, our final step in data preparation included aggregating individual data at the program level by calculating the mean score of each visitor outcome for each program. For SEM purposes, all data was then grand mean centered (Tabachnick & Fidell, 2007)

Dependent variables: outcomes

Based on extensive input from the NPS and a review of literature, we developed three dependent variables (Table 1). The first dependent variable served as a measure of visitor satisfaction with the program on a scale from 0 to 10, with 0=Terrible and 10=Excellent. Two indexes were developed from other survey items following procedures outlined by DeVellis (2003) to represent visitor experience and appreciation and behavioral intentions. The items comprising each index were measured using a five-point Likert-type scale, with answer choices: Not at all (1), A little (2), Somewhat (3), A moderate amount (4), and A great deal (5). Composite scores were created for each of the scales by taking the mean of all items (for more detail, see Stern & Powell, this issue).

Program and interpreter characteristics

The independent variables used in this SEM analyses included both interpreter and program characteristics that met two criteria. We included ordinal variables that were correlated ($p < 0.01$) to the particular outcome in question in any context (See Stern & Powell, article 1, and Powell and Stern, article 3, this issue). We also included categorical variables with at least "moderate" effect size in association with the particular outcome in question in any context (Cohen's $d > 0.5$). The program characteristics (Table 2) were originally drawn from an extensive literature review aimed at identifying best practices in the field (see Skibins et al., 2012). The interpreter characteristics were developed from the communications and education literature, though many of these factors are also referenced in the interpretation literature (Table 3). The tables also contain descriptive statistics. For more detail, see Stern and Powell (this issue).

Table 2. Program characteristics, their definitions, and descriptive statistics.

Program characteristic	Definition
Organization (α = 0.82) Scale: 1 to 5 Mean: 3.34 S.D.: 0.71	**Equally weighted composite mean score of 6 program characteristics:** • **Quality of the introduction** (Brochu and Merriman, 2002; Ham, 1992; Jacobson, 1999): Degree to which the introduction captured the audience's attention and oriented (or pre-disposed) the audience to the program's content and/or message. • **Appropriate sequence** (Beck and Cable, 2002; Ham, 1992; Jacobson, 1999; Larsen, 2003): Degree to which the program followed a logical sequence. • **Effective transitions** (Beck and Cable, 2002; Brochu and Merriman, 2002; Ham, 1992; Jacobson, 1999; Larsen, 2003): Degree to which program used appropriate transitions that kept the audience engaged and did not detract from the program's sequence. • **Holistic story** (Beck and Cable, 2002; Larsen, 2003; Tilden, 1957): Degree to which the program aimed to present a holistic story (with characters and a plot) as opposed to disconnected pieces of information. • **Clarity of theme** (Beck and Cable, 2002; Brochu and Merriman, 2002; Ham, 1992; Jacobson, 1999; Knudson, Cable, and Beck, 2003; Larsen, 2003; Lewis, 2005; Moscardo, 1999; Sharpe, 1976; Veverka, 1998; Ward and Wilkinson, 2006): 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. • **Link between introduction and conclusion** (Beck and Cable, 2002; Brochu and Merriman, 2002; Larsen, 2003): Degree to which program connected conclusion back to the introduction in an organized or cohesive way (i.e., program "came full circle.")
Connection (α = 0.88) Scale: 1 to 5 Mean: 2.77 S.D.: 0.78	**Equally weighted composite mean score of 5 program characteristics** • **Link tangibles to intangible meanings and universal concepts** (NPS Module 101; Beck and Cable, 2002; Brochu and Merriman, 2002; Ham, 1992; Knudson, et al., 2003; Larsen, 2003; Lewis, 2005; Moscardo, 1999; Tilden, 1957; Ward and Wilkinson, 2006): Communication connected tangible resources to intangibles and universal concepts. • Cognitive engagement (Knudson, et al., 2003; Moscardo, 1999; Sharpe, 1976; Tilden, 1957; Veverka, 1998): Degree to which the program cognitively engaged audience members in a participatory experience beyond simply listening; i.e. calls to imagine something, reflect, etc. • **Relevance to audience** (Beck and Cable, 2002; Brochu and Merriman, 2002; Ham, 1992; Jacobson, 1999; Knapp and Benton, 2004; Lewis, 2005; Moscardo, 1999; NPS Module 101; Sharpe, 1976; Tilden, 1957; Veverka, 1998): Degree to which the program explicitly communicated the relevance of the subject to the lives of the audience. • **Affective messaging** (Jacobson, 1999; Lewis, 2005; Madin and Fenton, 2004; Tilden, 1957; Ward and Wilkinson, 2006): Degree to which the program communicated emotion (in terms of quantity, not quality). • **Provocation** (Beck and Cable, 2002; Brochu and Merriman, 2002; Knudson, et al., 2003; Tilden, 1957): Degree to which the program explicitly provoked participants to personally reflect on content and its deeper meanings.
Appropriate logistics Scale: 1 to 4 Mean: 3.11 S.D.: 0.93	Degree to which basic audience and program needs were met (i.e., restrooms, weather, technology, accessibility, shade, etc). (Jacobson, 1999; Knudson et al., 2003)
Appropriate for audience Scale: 1 to 5 Mean: 3.93 S.D.: 0.70	Degree to which the program aligned with audience's ages, cultures, and level of knowledge, interest, and experience. (Beck and Cable, 2002; Jacobson, 1999; Knudson et al., 2003)
Multisensory Scale: 1 to 3 Mean: 2.39 S.D.: 0.51	Degree to which the program intentionally and actively engaged more than just basic sight and sound. (Beck and Cable, 2002; Knudson et al., 2003; Lewis, 2005; Moscardo, 1999; Tilden, 1957; Veverka, 1998; Ward and Wilkinson, 2006)
Physical engagement Scale: 1 to 4 Mean: 1.42 S.D.: 0.69	Degree to which the program physically engaged audience members in a participatory experience; i.e., through touching or interacting with resource. (Beck and Cable, 2002; Knudson, et al., 2003; Lewis, 2005; Moscardo, 1999; NPS Module 101; Sharpe, 1976; Tilden, 1957)
Verbal engagement Scale: 1 to 5 Mean: 2.51 S.D.: 1.02	Degree to which the program verbally engaged audience members in a participatory experience; i.e., dialogue (a two-way discussion). (Knudson, et al., 2003; Moscardo, 1999; Sharpe, 1976; Tilden, 1957; Veverka, 1998)
Fact-based messaging Binary: 27%	Program communicated only fact-based information. (Frauman and Norman, 2003; Jacobson, 1999; Lewis, 2005; Tilden, 1957; Ward and Wilkinson, 2006)
Clear message Scale: 1 to 4 Mean: 2.20 S.D.: 0.94	Degree to which program's message(s) was clearly communicated; i.e., the "so what?" element of the program. (Beck and Cable, 2002; Brochu and Merriman, 2002; Ham, 1992; Jacobson, 1999)
Consistency Scale: 1 to 3 Mean: 2.88 S.D.: 0.37	Degree to which the program's tone and quality were consistent throughout the program. (Beck and Cable, 2002; Ham, 1992)

Table 3. Interpreter characteristics observed in the study, their definitions, and descriptive statistics for cases analyzed in this paper.

Interpreter characteristic	Definition
Confidence (a = 0 .70) Scale: 1 to 4 Mean: 3.28 S.D.: 0.49	**Equally weighted composite mean score of 3 interpreter characteristics**: • **Comfort of the Interpreter** (Lewis 2005; Moscardo, 1999; Ward and Wilkinson, 2006): Degree to which the interpreter presenting the program seems comfortable with the audience and capable of successfully presenting the program without apparent signs of nervousness or self-doubt. • **Apparent knowledge** (Ham and Weiler, 2002a; Lewis, 2005; Ward and Wilkinson, 2006): The degree to which the interpreter appears to know the information involved in the program, the answers to visitors questions, and has local knowledge of the area and its resources. • **Eloquence** (Lewis, 2005): The extent to which the interpreter spoke clearly and articulately, and did not mumble or frequently use filler words such as "um" or "like."
Authentic emotion and charisma (a = 0.85) Scale: 1 to 5 Mean: 3.57 S.D.: 0.85	**Equally weighted composite mean score of 3 interpreter characteristics**: • **Passion** (Beck and Cable, 2002; Ham and Weiler, 2002b; Moscardo, 1999): The interpreter's apparent level of enthusiasm for the material, as opposed to a bored or apathetic attitude toward it. The overall vigor with which the material is presented. • **Charisma** (Ward and Wilkinson, 2006): A general sense of the overall likeability/charisma of the interpreter, commonly recognized by seemingly genuine interaction with the visitors, including smiling, looking people in the eye, and having an overall appealing presence. • **Sincerity** (Ham, 2009): The degree to which the interpreter seems genuinely invested in the messages he or she is communicating, as opposed to reciting information, and seems sincere in the emotional connection they may exude to the message and/or the resource. In other words, the extent to which the interpretation was delivered through authentic emotive communication.
Responsiveness Scale: 1 to 3 Mean: 2.81 S.D.: 0.41	The extent to which the interpreter interacts with the audience, collects information about their interests and backgrounds, and responds to their specific questions and requests or non-verbal cues. (Jacobson, 1999; Knudson et al., 2003; Lewis, 2005)
Humor quality Scale: 1 to 4 Mean: 2.08 S.D.: 0.73	How funny is the interpreter overall? Does the audience react positively to the interpreter's use of humor and seem to enjoy it? (Ham and Weiler, 2002b; Knapp and Yang, 2002; Regnier et al., 1992)
Sarcasm Scale: 1 to 3 Mean: 1.23 S.D.: 0.46	The degree to which the interpreter used sarcasm (the use of mocking, contemptuous, or ironic language or tone) or self-deprecation that was not meant to be serious, as a part of presenting their program.
Audibility Scale: 1 to 3 Mean: 2.86 S.D.: 0.36	The extent to which the interpreter could be clearly heard and understood by the audience.
Impatience Binary: 1.8%	Exhibition of explicit impatience toward audience members.
Goal: Behavior Change Binary: 7%	Intention of the interpreter for the program to influence audience's behavior. (Ham, 2013)

Table 4: Variables included in hypothesized models for each outcome.

Variable	Satisfaction	Visitor Experience and Appreciation	Behavioral Intentions
Interpreter characteristics			
Audibility	X	X	
Authentic emotion and charisma	X	X	X
Confidence	X	X	X
False assumption about audience	X	X	
Goal: Behavior change			X
Humor quality	X	X	X
Impatience	X		
Responsiveness	X	X	
Sarcasm			X
Program characteristics			
Appropriate for audience	X	X	X
Appropriate logistics	X	X	X
Clear message	X	X	X
Consistency	X	X	
Connection	X	X	
Multisensory engagement	X		
Organization	X	X	
Verbal engagement	X	X	X
Fact-based messaging	X		

Structural equation modeling

We used structural regression modeling (a.k.a. path analysis), a form of SEM, to examine the influence of different program and interpreter characteristics on three outcomes. We used SEM for this analysis because it is confirmatory (as opposed to exploratory) in nature and requires the researcher to have an explicit hypothesized model; it can model measurement error, which reduces inaccuracies; it allows for the analysis of a complete multivariate model including direct and indirect effects and in this case it can assess causal relationships between independent variables and a dependent variable (Byrne, 2006; Kline, 2005). In this study, all independent variables are formative (as opposed to reflective). That is, they were observed and represent a specific practice or attribute that is thought to directly influence the dependent variables (see Kline, 2005; Diamantopoulis et al., 2008; Diamantopoulis & Winklhofer, 2001; Jarvis et al., 2003, Padsakoff et al., 2007 for further explanation).

We used the EQS v6.1 software (Bentler 2005) to perform the statistical analyses, which progressed in several stages. First, the data were screened for univariate and multivariate deviations from normality. Next, we used structural regression modeling to assess the causal relationships between independent variables and each dependent variable (three separate models). For each outcome, we began with a model that contained all interpreter and program characteristics that met the criteria described

above for that outcome. The starting list of program practices and interpreter attributes used in the hypothesized models are in Table 4. To develop the final structural regression models we used an iterative process in which diagnostics (modification indices: Lagrange Multiplier Test (LM), Wald Test) indicated potential modifications, including removal of independent variables from the model, to improve fit and parsimony.

Structural regression analysis provides multiple statistics that can be used to evaluate the "fit" of a specified model (Byrne, 2006). In this paper we report the Satorra-Bentler Scaled Chi-Square (S-B χ^2), Robust Comparative Fit Index (CFI), Standardized Root Mean Square Residual (SRMR), the Robust Root Mean Square Error of Approximation (RMSEA) and its associated 90% confidence interval (Bentler & Yuan, 1999; Byrne, 2006). The S-B χ^2, which should be interpreted like a χ^2, is reported because it corrects for the degree of kurtosis in the data (Satorra & Bentler, 1994). The Robust CFI accounts for non-normality in the data and is an "incremental or comparative fit index" that evaluates the change in fit between the hypothesized model and the "independence model" (Byrne, 2006, 97; Bentler, 1990; Kline, 2005, 140). The independence model assumes that all the variables in the model are unrelated. The CFI represents the total covariation in the data and is measured on a scale of 0 to 1 with values greater than .9 indicating an acceptable fit and values greater than .95 indicating an excellent fit (Byrne, 2006; Hu & Bentler, 1999). The SRMR statistic provides the average difference between the sample and the predicted correlation matrices and thus is not susceptible to non-normality (Byrne, 2006). The SRMR uses standardized values with the range of scores between 0 and 1; values less than .1 are considered acceptable and less than .05 are considered a good fit (Hu & Bentler, 1995; Kline, 2005). The Robust RMSEA also accounts for non-normality in the data and is based on the average lack of fit per degree of freedom; therefore, as the fit improves, the RMSEA decreases. As such, this measure is sensitive to the degrees of freedom and the complexity of the model (Byrne, 2006). Like the SRMR, the scores range between 0 and 1, with values of .05 to .08 deemed acceptable and values less than .05 considered excellent (Browne & Cudeck, 1993; Hu & Bentler, 1999).

Beta weights in structural regression models reflect the effect size of an independent variable on the dependent variable. R^2 values gauge the predictive validity of the structural model, explaining the proportion of the total observed variance in the dependent variable explained by the model. It is recommended to assess R^2 values independently of fit indices, as the latter do not pertain to predictive validity (Kline, 2005).

Results

Three models were created based on the list of variables in Table 4. All independent variables (interpreter and program characteristics) were first entered as direct predictors of each outcome. In each case, the initial fit of each model was deemed unacceptable (Byrne, 2006). Through an iterative process, we adjusted the models using diagnostics that indicate potential model changes that would improve fit and parsimony. This generally involves removing variables one at a time based on statistical indicators produced at each stage of the modeling process. As the iterative modeling continues, it also can include adding or changing the nature of relationships between variables. In the end, a single "best fit" model is produced that represents the most parsimonious and predictive model for each outcome. The resulting models are displayed in Figures 1, 2, and 3.

Figure 1 represents the final model pertaining to how the interpreter and program

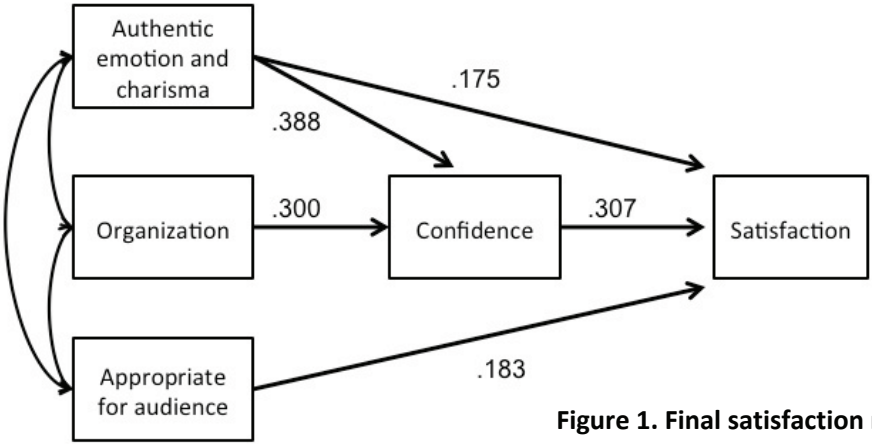

Figure 1. Final satisfaction model.

characteristics influenced *visitor satisfaction*. Fit indices for the final "satisfaction" model (SBχ^2=5.39, p < .07; CFI=.99; SRMR=.029; RMSEA=.08) indicated the model was an acceptable representation of the relationships present in the data (Byrne, 2006; Kline, 2005). *Authentic emotion* was a strong predictor of interpreter's confidence (β=.388, p < .05) and a weaker predictor of visitor *satisfaction* (β=.171, p < .05). Organization was also a strong predictor of interpreter's *confidence* (β=.300, p < .05), but not a direct predictor of visitor *satisfaction*. Confidence was a strong predictor of visitor *satisfaction* (β=.307, p < .05). *Appropriate for the audience* was also a significant predictor of visitor *satisfaction* (β=.183, p < .05). The model accounted for 35% (R^2) of the variance in confidence and 27% (R2) of the variance in visitor *satisfaction*.

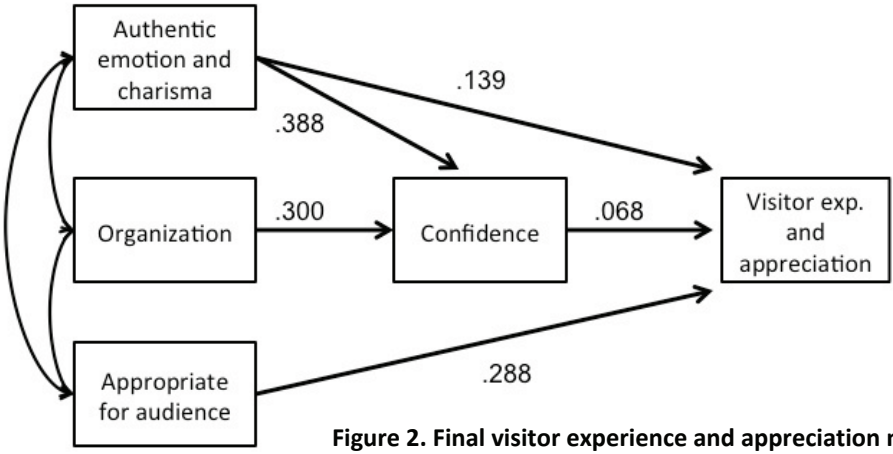

Figure 2. Final visitor experience and appreciation model.

The final structural regression model for visitor *experience and appreciation* had the same structure as the final visitor satisfaction model (Figure 2). Fit indices for the model (SBχ^2=4.45, p < .1; CFI=.99; SRMR=.027; RMSEA=.069) indicated the model was an acceptable fit of the data (Byrne, 2006; Kline, 2005). The only structural differences between this model and the satisfaction model involved the relative strength

of *confidence* and *appropriateness for the audience*. *Appropriate for audience* (β=.288, p < .05) was the strongest predictor of visitor experience and appreciation, followed by *authentic emotion* (β=.139, p < .05) and *confidence* (β=.068, p < .05). The model accounted for 35% (R^2) of the variance in *confidence* and 16% (R^2) of the variance in *visitor experience and appreciation*.

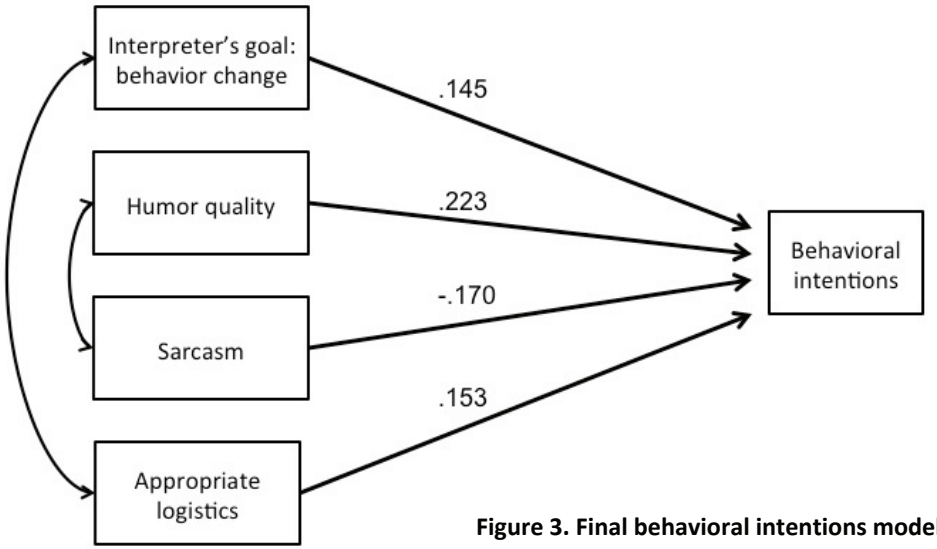

Figure 3. Final behavioral intentions model.

The model in Figure 3 represents how interpreter and program characteristics predicted *intentions to change behaviors*. Fit indices for the model in Figure 3 ($SB\chi^2$=7.38, p < .05; CFI=.96; SRMR=.040; RMSEA=.03) indicated the model was an acceptable representation of the relationships present in the data. Having a *goal to influence behavior* (β=.145, p < .05), *appropriate logistics* (β=.153, p < .05), and *humor quality* (β=.223, p < .05) were significant positive predictors of intentions to *change behaviors*. Use of *sarcasm* (β=-.170, p < .05) was a significant but negative predictor of *intentions to change behaviors*. The model accounted for 10% (R^2) of the variance in intentions to change behaviors.

Discussion: Is it the interpreter or the program?

We used structural equation modeling to examine the relative influence of interpreter and program characteristics upon visitor outcomes at live interpretation programs across the U.S. National Park Service. The resulting models reveal three main lessons. First, it appears in each case that *both* interpreter and program characteristics influenced visitor outcomes. Second, depending on outcome, certain program practices and interpreter attributes provided the best model fit and predictive power. Third, the final models accounted for a relatively low percentage of the overall variance in visitor outcomes. We explain each finding and some important limitations in the interpretation of the analyses below.

In each model, both interpreter and program characteristics influenced outcomes. The *satisfaction* and the *visitor experience and appreciation* models each contained *authentic emotion and charisma, organization, confidence,* and *appropriate for the*

audience. In each model, *authentic emotion and charisma* and *organization* were mediated by *confidence*. In other words, the model suggests that authentic emotion and charisma and organization each help to create interpreter confidence, which in turn enhances visitor outcomes. *Authentic emotion and charisma* also served as a direct causal predictor of each outcome, as did the *appropriate for the audience variable*.

The final structural regression model of intentions to change behaviors suggests that *humor quality, appropriate logistics,* and *intending to influence* behaviors through a program positively influenced intentions to change stewardship behaviors. The use of *sarcasm* was associated with weaker intentions to change stewardship behaviors. In other words, interpreters that successfully employed *humor*, ensured that their audience's needs were met, and explicitly intended to influence their audience's behaviors were more successful at doing so. Meanwhile, overly sarcastic interpreters were less likely to influence changes in behavioral intentions. Interestingly, only 7% of all interpreters interviewed in the study explicitly intended to influence audience behaviors (Table 3). Ham (2013) reminds interpreters that outcomes, such as behavior change, do not happen magically; instead a program should be planned and developed with an outcome in mind. When focusing on behavior change, numerous techniques may increase the likelihood of influencing specific behaviors (Ham et al., 2007; Powell & Ham, 2008; Stern & Powell, this issue).

Certain limitations in the data and analyses are important to consider when interpreting these findings. First, structural equation modeling explicitly aims to produce the most parsimonious predictive or in this case causal model for selected outcomes. As such, independent variables that may be strongly related to outcomes are commonly removed during the modeling process due to their relationships with other variables. For example, the connection variable is highly correlated with organization, authentic emotion and charisma, and confidence (see Stern et al., this issue). As a result, it may be removed from a model because it explains a redundant proportion of the variance in the outcome as the other independent variables. This is the case with many of the program practices and interpreter characteristics tested in this analysis. It would be inappropriate to assume that their absence in the final models reduces their significance in influencing more positive outcomes.

Second, the models accounted for 10% to 27% of the variance in the outcomes. The strongest model accounted for 27% of the variance in *satisfaction*. The weakest model accounted for 10% of the variance in *behavioral intentions*. This suggests that much more is at play than simply the interpreter and the program elements. Interpretive programs are complex phenomena, and audience outcomes can be influenced by characteristics of the individual audience members, the makeup of the group, and the location and context of the program, in addition to characteristics of the program and the interpreter (Powell et al., 2009). Past research into communications (see Ajzen, 1992, for more) suggests that few consistent trends emerge when attempting to examine the range of source (interpreter) factors, receiver (audience) factors, channel (program) factors, and message (content) factors that influence outcomes resulting from communications. These factors vary with each program and produce an almost unlimited number of interactions and potential combinations (Falk, 2004). We examine a small portion of these additional factors in a separate article in this issue (Powell & Stern, this issue).

Relatively low R^2 values may also be a product of the lack of variance observed in *satisfaction* and *visitor experience and appreciation* scores. We discuss this issue in

greater depth in a separate article in this issue as well (Stern et al., this issue). Predictive ability may be particularly low for *behavioral change* for a number of reasons. As noted earlier, few programs actually targeted behavioral change as an outcome. As such, changing behavioral intentions may have been more of a side effect than an intended outcome of a program. Moreover, many interpretive program goers may already perform many of the behaviors discussed in interpretive programs, leaving little room for change (see Stern & Powell, this issue, for a more detailed discussion).

Despite the limitations, the results suggest that outcomes are influenced by both program *and* interpreter characteristics and that these characteristics interact and influence each other. For example, confidence may ultimately emerge from an interpreter's passion for the resource and careful planning, which leads to good organization. Because most prior research and formal training have focused on what we have categorized as "program characteristics" (Skibins et al., 2012), we urge future researchers, trainers, and practitioners to give some meaningful attention to interpreter attributes and delivery styles. Training programs might add elements that could improve interpreters' abilities to project confidence and authentic emotion. Some lessons for doing so might be found in the formal education field, where "affinity-seeking" and immediacy behaviors have garnered some attention (e.g., Finn et al., 2009). These practices involve efforts to ingratiate teachers with their students by reducing the social distance between them (see also Stern & Powell, this issue; Stern et al., in press). Interpretive organizations might also consider these findings in light of the role of the individual interpreter in program development. If organizations can provide opportunities for creating and sustaining authentic connections between interpreters and the resources they interpret, they might enhance interpreters' abilities to convey their own passions to their audiences. Finally, we urge researchers to consider how different program and interpreter characteristics may function differently in varying contexts.

References

Ajzen, I. (1992). Persuasive communication theory in social psychology: A historical perspective. In M. J. Manfredo (Ed.), *Influencing human behavior* (pp. 1–27). Champaign, IL: Sagamore Publishing.

Archer, D., & Wearing, S. 2003. Self, space, and interpretive experience: The interactionism of environmental interpretation. *Journal of Interpretation Research,* 8 (1): 7–23.

Beck, L., & Cable, T. T. (2002). *Interpretation for the 21st century: Fifteen guiding principles for interpreting nature and culture* (2nd ed.). Champaign: Sagamore.

Bentler, P.M. 1990. Comparative fit indexes in structural models. *Psychological Bulletin,* 107: 238–46.

Bentler, P.M. 2005. *Eqs 6 structural equations program manual*. Encino, CA: Multivariate Software (www.mvsoft.com).

Bentler, P.M. and K.H. Yuan. (1999). Structural equation modeling with small samples: Test statistics. *Multivariate Behavioral Research,* 34: 181–97.

Brochu, L., & Merriman, T. (2002). *Personal interpretation: Connecting your audience to heritage resources*: Fort Collins, CO: InterpPress.

Browne, M.W., & Cudeck, R. (1993). *Alternative ways of assessing model fit. In Testing structural equation models*, ed. K.A. Bollen and J.S. Long, 445–55. Newbury Park, CA: Sage.

Byrne, B.M. (2006). *Structural equation modeling with eqs: Basic concepts, applications and programming.* Second ed. Mahwah, NJ: Erlbaum.

DeVellis, R.F. (2003). *Scale development: Theory and applications Applied social research methods*. 2nd ed. Thousand Oaks, CA: Sage Publishing.

Diamantopoulos, A., Riefler, P., & Roth, K. P. (2008). Advancing formative measurement models. *Journal of Business Research*, 61(12), 1203–1218.

Diamantopoulos, A., & Winklhofer, H.M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research*, 38(2); 269–277

Falk, J. (2004). The director's cut: Toward an improved understanding of learning from museums. *Science Education*, 88(S1), S83–S96. doi: 10.1002/sce.20014

Finn, A. N., Schrodt, P., Witt, P. L., Elledge, N., Jernberg, K. A., & Larson, L. M. (2009). A meta-analytical review of teacher credibility and its associations with teacher behaviors and student outcomes. *Communication Education*, 58(4), 516–537.

Forist, B. (2003). *Visitor Use and Evaluation of Interpretive Media*. A Report on Visitors to the National Park System. National Park Service Visitor Services Project. http://nature.nps.gov/socialscience/ docs/Visitor_Use_and_Evaluation.pdf. Accessed May 24, 2012.

Fox, J. 1991. *Regression diagnostics*. Newbury Park, CA: Sage.

Frauman, E., & Norman, W.C. (2003). Managing visitors via "mindful" information services: One approach in addressing sustainability. *Journal of Park and Recreation Administration*, 21 (4): 87–104.

Ham, S. (2009). From interpretation to protection: Is there a theoretical basis? *Journal of Interpretation Research* 14(2): 49–57.

Ham, S. H. (1992). *Environmental interpretation: A practical guide for people with big ideas and small budgets*. Golden, CO: Fulcrum Publishing.

Ham, S. H. (2013). *Interpretation: Making a difference on purpose*. Golden, CO: Fulcrum Publishing.

Ham, S. H., Brown, T. J., Curtis, J., Weiler, B., Hughes, M., & Poll, M. (2007). *Promoting persuasion in protected areas: A guide for managers*. Developing strategic communication to influence visitor behavior. Southport, Queensland, Australia: Sustainable Tourism Cooperative Research Centre.

Ham, S.H., & Weiler, B.M. (2002a). Toward a theory of quality in cruise-based nature guiding. *Journal of Interpretation Research,* 7 (2): 29–49.

Ham, S H., & Weiler, B.M. (2002b). Tour guide training: a model for sustainable capacity building in developing countries. *Journal of Sustainable Tourism,* 10(1): 52–69.

Hu, L.-T., & P.M. Bentler. (1995). Evaluating model fit. In *Structural equation modeling: Concepts, issues, and applications*, ed. R. H. Hoyle, 76–99. Thousand Oaks, CA: Sage.

Hu, L.-T., & P.M. Bentler. (1999). Cutoff criteria for fit indices in covariance structure analysis: Guidelines, issues, and alternatives. *Structural Equation Modeling,* 6: 1–55.

Jacobson, S. K. (1999). *Communication skills for conservation professionals*. Washington, D.C.: Island Press.

Jarvis, C.B., MacKenzie, S.B., & Podsakoff, P.M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, 30(2); 199–218.

Kline, R.B. (2005). *Principles and practice of structural equation modeling.* 2nd ed. New York: The Guilford Press.

Knapp, D., & Benton, G. M. (2004). Elements to successful interpretation: A multiple case study of five national parks. *Journal of Interpretation Research*, 9(2), 9–25.

Knapp, D. & Yang, L. (2002). A phenomenological analysis of long-term recollections of an interpretive program. *Journal of Interpretation Research*, 7(2), 7–17.

Knudson, D. M., Cable, T. T., & Beck, L. (2003). *Interpretation of cultural and natural resources,* (2nd ed.). State College: Venture Publishing.

Larsen, D. L. (2003). *Meaningful Interpretation: How to Connect Hearts and Minds to Places, Objects, and Other Resources*. Fort Washington: Eastern National.

Lewis, W. J. (2005). *Interpreting for park visitors* (9th ed.). Fort Washington: Eastern National.

Madin, E. M. P., & Fenton, D. M. (2004). Environmental interpretation in the Great Barrier Reef Marine Park: An assessment of programme effectiveness. *Journal of Sustainable Tourism*, 12(2), 121–137.

McCroskey, J.C., Richmond, V.P., & Stewart, R.A. (1986). *One on one: The foundations of interpersonal communication*. Englewood Cliffs, NJ: Prentice-Hall.

McManus, P.M. (1987). It's the company you keep . . . The social determination of learning-related behaviour in a science museum. *The International Journal of Museum management and Curatorship,* 6: 263–270.

McManus, P. M. (1988). Good companions: More on the social determination of learning-related behaviour in a science museum. *International Journal of Museum Management and Curatorship*, 7(1), 37–44.

Moscardo, G. (1999). *Making visitors mindful: Principles for creating quality sustainable visitor experiences through effective communication*. Champaign: Sagamore.

National Park Service. (2003a). *Interpretive Development Program. Module 101: Fulfilling the NPS mission: The process of interpretation*.

National Park Service. (2003b). *Interpretive Development Program. Module 103: Preparing and presenting an effective interpretive talk*.

National Park Service. (2003c). *Interpretive Development Program. Module 210: Prepare and present an effective conducted activity.*

National Park Service. (2003d). *Interpretive Development Program. Module 220: Prepare and present an interpretive demonstration or other illustrated program*.

National Park Service. (2003e). *Interpretive Development Program. Module 230: Interpretive writing.*

National Park Service. (2003f). *Interpretive Development Program. Module 311: Interpretive media development*.

Podsakoff, P.M., MacKenzie, S.B., Lee, J.Y., & Podsakoff, N.P. (2007). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5) 879–903.

Powell, R.B., Brownlee, M.T.J., Kellert, S. R. & Ham, S.H. (2012) From awe to satisfaction: Immediate affective responses to the Antarctic tourism experience. *Polar Record*. 48(2): 145–156.

Powell, R.B., Kellert, S. R., & Ham, S.H. (2009). Interactional theory and the sustainable nature-based tourism experience. *Society and Natural Resources*, 22(8): 761–776.

Powell, R. B., & Ham, S. H. (2008). Can ecotourism interpretation really lead to pro-conservation knowledge, attitudes and behavior? Evidence from the Galapagos Islands. *Journal of Sustainable Tourism*, 16(4): 467–489.

Regnier, K., Gross, M., & Zimmerman, R. (1992). *The Interpreter's Guidebook: Techniques for Programs and Presentations*. Interpreter's Handbook Series. Stevens Point: UW-SP Foundation Press.

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors on covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis* (pp. 399–419). Thousand Oaks, CA: Sage.

Skibins, J.C., Powell, R.B., & Stern, M.J. (2012). Exploring empirical support for interpretation's best practices. *Journal of Interpretation Research*. 17(1): 25–44.

Sharpe, G. W. (1976). *Interpreting the environment*. New York: John Wiley & Sons.

Stokols, D., & Altman,I. eds. (1987). *Handbook of environmental psychology*. New York: John Wiley & Sons.

Tabachnick, B.G. and L.S. Fidell. (2007). *Using multivariate statistics* (5th ed.). Needham Heights, MA: Allyn & Bacon.

Tilden, F. (1957). *Interpreting our heritage* (3rd ed.). Chapel Hill: The University of North Carolina Press.

Veverka, J. A. (1998). *Interpretive master planning: the essential planning guide for interpretive centers, parks, self guided trails, historic sites, zoos, exhibits and programs* (2nd ed.). Tustin: Acorn Naturalists.

Wearing, S., & Wearing, B. (2001). Conceptualizing the selves of tourism. *Leisure Studies*, 7: 11–23.

Ward, C.W., & Wilkinson, A. E. (2006). *Conducting meaningful interpretation: A field guide for success*. Golden: Fulcrum.

Copyright of Journal of Interpretation Research is the property of National Association for Interpretation and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.