

2009

An assessment of the outcomes of a Chinese-language interpretive tour experience at a heritage tourism attraction

Pieter Van Dijk
Monash University

Betty Weiler
Monash University

Publication details

Postprint of: Van Dijk, P & Weiler, B, 2009, 'An assessment of the outcomes of a Chinese-language interpretive tour experience at a heritage tourism attraction', *Tourism Analysis*, vol. 14, no. 1, pp. 49-63.

Published version available from: <http://dx.doi.org/10.3727/108354209788970144>

ePublications@SCU is an electronic repository administered by Southern Cross University Library. Its goal is to capture and preserve the intellectual output of Southern Cross University authors and researchers, and to increase visibility and impact through open access to researchers around the world. For further information please contact epubs@scu.edu.au.

Title:

An Assessment of the Outcomes of a Chinese-language Interpretive Tour Experience at a Heritage Tourism Attraction

Authors:

Pieter van Dijk, Tourism Research Unit, Monash University

Betty Weiler, Tourism Research Unit, Monash University

Accepted on 07/10/2008 for *Tourism Analysis* (volume and issue TBD)

Acknowledgements

The authors wish to thank Dr Xin Yu and Andrea Zhang for assistance with translation and back-translation of the instrument. Andrea was also the main research assistant for data collection and her dedication and good humour were crucial in gaining a high rate of response from visitors. We also thank Leah Stamm for her help with data entry and analysis and Annita Allman for assisting with data and graphic presentation. The authors are particularly grateful to Tim Sullivan, Deputy CEO with The Sovereign Hill Museums Association, for his enthusiasm and support in undertaking the research.

An Assessment of the Outcomes of a Chinese-language Interpretive Tour Experience at a Heritage Tourism Attraction

Abstract

This study uses the first translated version of an established and validated instrument to assess the impacts of an interpretive tour experience on Chinese visitors to a heritage tourism attraction in Australia. The study makes a significant methodological contribution by way of testing the reliability of this instrument in a Chinese cultural context. In addition to its methodological contribution, the study represents the first known published research evaluating the impacts of Chinese-language interpretive tours on inbound visitors from China. Based on results from 285 self-completed surveys, the instrument reliably assessed that visitors' perceptions of their interpretive experience were average to very good on cognitive, affective and behavioral constructs. Notwithstanding the fact that work still needs to continue in refining one of the ten indicators, the instrument provides practitioners and researchers with a reliable and valid tool for assessing the outcomes of heritage interpretation in two very different cultural contexts.

Key words: Chinese tour groups; heritage tourism attractions; interpretation; visitor outcomes

Introduction

Tourism operators of all kinds seek to use on-site communication by tour guides, interpreters and other customer service staff to enhance the visitor experience, ostensibly because such communication is linked to visitor satisfaction, positive word-of-mouth advertising and repeat visitation (Wearing et al., 2006; Ham & Weiler, 2007). In addition, interpretation is becoming much more widely understood and used by tourism attractions and tour operators as an important tool for promoting a range of other cognitive, affective and behavioral visitor outcomes (de Rojas & Camarero, 2008; Weiler & Ham, 2006). However, evaluation methods and tools that can gauge the success of communication in achieving these outcomes have been lacking or, at best, unreliable, especially for cross-site and non-English (e.g. Chinese language) contexts.

The aims of this study were to develop and test a Chinese version of a validated interpretation evaluation instrument and to use it to assess the impacts of an interpretive tour experience on Chinese-speaking tour groups visiting Sovereign Hill, a re-created gold-rush heritage attraction in Australia. In particular, the study set out firstly to examine the reliability of a translated version of an interpretation evaluation instrument for assessing outcomes, and secondly to measure the cognitive, affective and behavioral outcomes reported by visitors as a result of their experience. It is not the aim of the study to make comparisons between Chinese and non-Chinese visitors, but to report the methods and findings as being of importance in their own right and of interest to any country where China is (or soon will be) one of the top source markets. In this paper, we use the terms interpretation, interpretive experience, and on-site communication interchangeably to refer to the Chinese-language interpretation provided by the on-site guides.

The paper begins with a review of literature on heritage tourism experiences with a focus on previous research concerning the impacts of visitor communication, in particular heritage interpretation, and on what visitors take away from those experiences. This is followed by consideration of a range of methodological issues in assessing the impacts of interpretation on visitors and on the experience, drawing on both program evaluation methods generally and the evaluation of interpretation in particular, which together served as a basis for the selection and refinement of the method and instrument used in this study. The remainder of the paper presents the context for the empirical study, the rationale for the selection of the study population and site, a discussion of the study methods and the findings. Firstly, results in relation to reliability-testing of the research instrument in a Chinese language context are outlined. Secondly, results are presented with respect to the extent to which visitors perceived that a range of cognitive (thought), affective (attitudes and emotions) and behavioral (intentions to act) outcomes occurred as a result of their experience. The paper concludes with implications for the effective use of interpretation as a tool for enhancing the experiences of Chinese visitors and the outcomes of that experience, together with a synopsis of the study's contribution to the measurement and evaluation of visitor outcomes, including an evaluation tool adapted for and tested in a Chinese language context.

Literature Review

While interpretation is perhaps best known for its use by public sector attractions such as museums and national parks, in recent years heritage tourism managers and operators have become more aware that visitors expect interpretation and believe it is an important part of their experience (Morgan et al., 2006). Interpretive products such as guided walks, tours, interpretive talks, demonstrations and other kinds of face-to-face interpretive activities as well as many static (non-personal) media are increasingly used to enhance the visitor experience and thereby add a competitive advantage as well as to assist with other strategic outcomes.

These outcomes include engaging visitors, increasing customer loyalty, promoting good public relations, fostering positive attitudes toward conservation, managing visitors' on-site behaviors for a range of purposes such as minimizing visitor risk and adverse environmental impacts, and even influencing post-visit, long-term behavior (Knudson et al., 1995). As a result, there has been increasing interest in research that measures and evaluates the effectiveness and outcomes of interpretation (Ballantyne and Uzzell, 1999). However, widespread agreement on how best to evaluate the success of visitor communication/interpretation is lacking. Thus it was important to review both generic program evaluation literature and interpretation evaluation studies to identify theoretical and methodological advancements that could inform the research.

In the first instance, successful communication with visitors needs to be acknowledged as that which achieves its purposes, and thus there needs to be consideration of the purposes of visitor communication in order to gauge its success. So, while there is certainly merit in evaluating interpretation against management inputs such as its cost, there is growing recognition of the value of being able to evaluate interpretation in relation to its outputs, *i.e.* its impact on visitors. Henderson & Bialeschki (1995) propose that “determining the impact, effects, outcomes, and results of a program, area or facility, or administrative procedure is the bottom line of evaluation” (p. 11). Thus, it can be argued that evaluating what visitors take away as a result of an interpretive experience is critical to determining its success.

In a review of interpretation and evaluation research that spanned more than two decades and focused on identifying and measuring visitor outcomes in a wide range of contexts and settings, interpretation was found to potentially impact three overarching constructs or domains of interest: visitor cognition (what visitors think), affect (what visitors feel) and behavior (what visitors do) (Ham & Weiler, 2006). These correspond to Ham's (1982)

classification of the outcomes of interpretation of some 25 years ago: (1) cognitive responses (visitor learning, knowledge acquisition, or information processing), (2) affective responses (visitor feelings, attitudes or emotions), and (3) behavioral responses (visitor actions or inactions). While others have used slightly different classification systems (e.g. Dierking *et al.*, 2004; Wearing *et al.*, 2006), there is consistency in terms of the constructs captured. Therefore, this classification of domains of interest – cognition, affect and behavior – is used in the following discussion to present a review of interpretive research at museums, heritage sites and attractions, and in particular to consider studies that have assessed one or more of these indicators in the context of experiences at heritage sites and attractions.

There has been a considerable amount of research evaluating interpretation in museums and some interpretive evaluation at heritage sites and attractions (Uzzell, 1998a; Uzzell, 1998b) but the vast majority of this research has focused on static, non-personal interpretive media such as museum displays and exhibits. Also, rather than focusing on what visitors take away from these experiences, many of these studies have been primarily concerned with other on-site observable variables such as attention paying, viewing time, holding power, and reading behavior (*cf.* Bicknell & Mann, 1993; Bitgood *et al.*, 1988; Bitgood & Patterson, 1993; Boisvert & Slez, 1994; Camp *et al.*, 2000; Chiozzi & Andreotti, 2001; Fernandez & Benlloch, 2000; Harvey *et al.*, 1998; Korn & Jones, 2000). Several museum researchers have looked at duration and content of conversation as surrogate indicators of visitor outcomes (Ash, 2003; Fernandez & Benlloch, 2000; Leinhardt & Crowley, 1998; Leinhardt & Knutson, 2004; Leinhardt *et al.*, 2002; McManus, 1987; McManus, 1989) but again have not attempted to demonstrate empirically the links between these variables and visitor cognition, affect or behavior. Some researchers, for example Beeho & Prentice (1995), Light (1995) and Prentice *et al.* (1998), have sought to describe the on-site behavior and experience of the visitor and to some extent what visitors have learned. Beyond learning, which is largely measured by using

recall, there had been very limited heritage/museum research prior to the early 1990s that measured or quantified the outcomes of interpretation.

In the past two decades, a growing body of research has emerged looking beyond visitor recall to other cognitive, affective and/or behavioral outcomes of museum experiences, for example, Allard *et al.* (1994); Allen (1997); Anderson *et al.* (2000); Blud (1990); Cartmill & Day (1997); D'Agostino *et al.* (1992); White & Jacobson (1994); Falk (1997); Finson & Enochs (1987); Henriksen (2001); Hernandez & Celis Herrero (1993); Stevens & Hall (1997); Stevenson (1991); and Zeidler & Surber (1999). Most of these studies focus on one or at most two of the three domains of interest, and often using a fairly targeted range of indicators, often (and rightly so) related to organizational objectives and/or the particular interests of the researcher. As a result, comparisons and generalizations about the outcomes of interpretation have been constrained by the one-off, cross-sectional and site-specific nature of the research.

Toward the end of the decade, Stewart *et al.* (1998) were calling for “a more theory-driven approach to the evaluation of interpretation” (p.265). Also, Ham and Weiler (2006) noted that, in the absence of comparable study designs, methods and measurement tools, empirical evidence of the outcomes of interpretation in museums and heritage settings has been inconclusive and largely not generalizable. It was partly in response to this need for a theory-driven tool with which to benchmark, monitor and compare the outcomes of interpretation between visitor groups, across sites and over time, and between different types of tourism operations such as attractions, tours, heritage sites, protected areas, zoos, events and so on, that Ham and Weiler (2006) set out to produce their *Interpretation Evaluation Tool Kit*. More precisely, they undertook a process to develop, field-test, refine and deliver an industry-driven and theoretically-grounded set of methods and data collection instruments for assessing the outcomes of face-to-face interpretation in a range of settings including heritage sites and

attractions. The tool kit comes in three different forms aimed at evaluating food and beverage-based, nature-based and heritage-based interpretation outcomes. The heritage interpretation instrument developed as part of the tool kit assesses visitor outcomes on a set of indicators that: (1) reflects the types of outcomes that managers actually want from their interpretive programs (for example, enhanced visitor enjoyment, positive visitor attitudes toward conservation, positive word-of-mouth advertising, provoking visitors to think about the values inherent at the site, etc.), (2) is theoretically valid based on what is known about interpretation's potential impacts on visitor cognition, affect and behavior, and (3) requires minimal effort from respondents, yet produces results that are both valid and reliable. Further information about the development and validation process of the tool kit is available in Ham and Weiler (2006).

The heritage version of the data collection instrument (questionnaire) developed for the tool kit was selected for the present study because it offered a reliable, valid, theoretically-driven and field-tested assessment tool as well as two other distinct advantages. First, it provided a way of assessing the outcomes of interpretation that could serve as a benchmark for the attraction, for other sites and for other visitor groups. Second, the study sought to ascertain whether the instrument would hold up in a cultural context different to the one in which it was developed. Notably, the instrument was developed and tested only on English-speaking visitors and, to this end, the current study undertook to translate and test the instrument for use on Mandarin-speaking visitors for reasons that will be explained in the following section.

Study Population and Site

The remainder of this paper reports on an empirical study of inbound Chinese tourists visiting Australia as part of a tour group. Australia and New Zealand were the first Western

destinations to be granted Approved Destination Status (ADS) by the Chinese Government in 1999. Since then around 14,000 ADS groups have travelled to Australia.

Inbound travel by Chinese residents continues to grow rapidly and China is predicted to become the largest potential growth market not only for Australia (Department of Industry, Tourism and Resources [DITR], 2005) but also for other countries including those in Europe (China National Tourism Administration [CNTA], 2006, cited in Tourism Australia [TA], 2006) and Arrivals from China to Australia are predicted to increase by an average annual growth rate of more than 16 per cent over the next ten years (DITR, 2007; Weiler & Yu, 2008). The present study was motivated in part by the extraordinary growth in this market segment.

Under the ADS scheme, Chinese holidaymakers wishing to visit Australia must join a fully-inclusive ADS tour group, accompanied by a tour leader or guide from China. Most tour group members are relatively inexperienced international travellers, speak little or no English, and have worked long and hard for the chance to travel overseas. Typical ADS tours include the 12-day Australia and New Zealand tour and the 8-day Australia tour, the latter being popular among Chinese consumers in the Chinese New Year holiday (Weiler & Yu, 2008).

For the present study, Chinese tour group members who were visiting one particular heritage tourism attraction – Sovereign Hill – were surveyed during one of these Chinese New Year periods. Sovereign Hill is operated by Sovereign Hill Museums Association and is one of Australia's major tourist attractions, attracting almost half a million visitors per year. Located 110 km northwest of Melbourne in the regional urban centre of Ballarat, the trip to Sovereign Hill from Melbourne's CBD takes approximately 90 minutes one-way by coach. The re-created 1850s goldfields township with its costumed and highly skilled interpretive staff,

entertaining street theatre and engaging underground mine tours is known for its commitment to authentically depicting life on the goldfields from a number of perspectives, including the life of the estimated 9000 gold-diggers and miners who came in search of gold from southern China. The relevance of these latter stories is one of the reasons why Chinese tour wholesalers include Sovereign Hill in their very limited Australian itinerary.

As a result of the relatively short time that these visitors spend in and around Melbourne (typically only one or two nights), Chinese tour group visits to the Sovereign Hill attraction are often less than two hours compared to the half-day or full-day visit that is typical of Australians and other overseas visitors. Consequently, the areas within the site typically visited by Chinese tour groups represent only a subset of a much larger attraction and potential experience. Because the site has virtually no visitor navigational aids (for reasons of preserving a sense of authenticity) and because Chinese visitors for the most part cannot speak or understand English, they are almost completely dependent on where their Mandarin-speaking local guide takes them and what s/he tells them.

On the positive side, all groups are provided with bilingual local guides who do orient and interpret selected areas of the attraction. Sovereign Hill employs up to twelve casual Mandarin and Cantonese-speaking tour guides to cater for visitors from China and ethnic Chinese visitors from other Southeast Asian countries. A typical Chinese tour group is escorted by one of the Chinese-speaking tour guides who deliver commentary at the various sites. In the Underground Mine, visitors view a Chinese-language movie about the Gold Rush including the experience of Chinese gold-diggers. Aside from this audio-visual presentation, visitors are reliant on the Chinese-speaking tour guide to interpret the history of gold-mining

in Australia and life on the goldfields. Depending on the guide, the commentary can be largely factual or it can be enlivened by stories and interesting anecdotes.

While several studies have examined tour-guiding and visitor interpretation in Taiwan (JuiChi, 2006; ShiuNan et al., 2005) there appears to be no published quantitative research on the outcomes of interpretation for Chinese-speaking visitors. Weiler and Yu (2007), however, provide a detailed qualitative report on the experience of Chinese visitors to Sovereign Hill which suggests that expectations of Chinese tourists are often high. Methodological advances in accurately assessing visitor outcomes associated with this market segment are therefore needed in order to effectively evaluate and improve tourism-based products designed for visitors from China. The following section outlines the methods applied for the present study.

Instrument Development

The *Interpretation Evaluation Tool Kit* is designed in consultation with industry, informed by previous academic research and further developed and refined using field testing and statistical validation. In developing the tool kit a structured facilitation process (modified Nominal Group Process) was conducted with multiple levels of staff at two sites (one being Sovereign Hill) that offered a number of face-to-face interpretation programs. The aim of this process was to determine the perceived desired outcomes of interpretation. During this process staff were required to complete the sentence: “At the end of the day, I would know if the ...interpretive program is doing a good job if...” Once a list of potential outcomes was created (e.g. “visitors enjoyed themselves”) the staff were then required to rank them. Following the ranking of outcomes, and due to overlaps and duplications between lists, three subject matter experts independently assigned responses to categories informed by human behavior research. If consensus on items could not be reached by the three subject matter experts they were deleted. This process reduced the original 54 items to 37. Indicators were

further reduced after inter-site comparisons were made and duplications eliminated. The final suite of indicators focused on three main categories of outcomes: cognitive, behavioral, and affective. In other words, the heritage questionnaire used for the present study was developed using a rigorous best-practice methodological approach that fits the requirements for development of a valid and reliable scale (*cf.* Ham & Weiler, 2006). In the final instrument, respondents' perceptions, or self-reports, of the impact of interpretation are assessed by way of 30 individual response items, most of which are seven-point scale items that together measure ten "indicators" or outcomes of interpretation in the three categories of cognition, affect and behavior. Table 1 presents the suite of indicators and items used in the questionnaire in relation to these three outcome domains together with their respective Cronbach's alpha coefficients. Further description of these indicators is included in the discussion of the results.

Insert Table 1 here

As the tours were conducted in Mandarin and respondents in this study for the most part not only could not speak and understand English but also could not read and write English, the first stage of the research was to translate the questionnaire into Mandarin. The procedure for translating and back translating the instrument followed recognised procedures (*cf.* Behling & Law, 2000; Brislin, 1970; Triandis & Brislin, 1984). Two bilingual (Mandarin and English) research assistants (RAs) were employed to carry out the translation. In the first step of the conversion process, the first RA translated the questionnaire into Mandarin. The second RA back-translated the questionnaire into English based on the initial RA's work. Back-translation is a recognized method used in order to preserve uniformity between the meaning of the translated and untranslated versions of an instrument (Behling & Law, 2000). Appropriate refinements were made after consultation with, and between, the RAs in order to

retain the integrity of the intended meaning of the questions. The instrument was then pilot-tested by one of the RAs and an additional Mandarin-speaking assistant on a respondent sample similar to the target population for this study. The feedback was evaluated and consequently some minor adjustments to the translated version of the instrument were made. The refined Mandarin version of the instrument was then used at Sovereign Hill with visiting tour groups. The final self-completed survey instrument took visitors between five and eight minutes to complete.

Field Methods and Limitations

The data for the study were collected over a five day period in February 2007, which included the peak period of Chinese New Year. Mandarin-speaking RAs invited visitors to complete the survey at the conclusion of their tour and prior to undertaking any unguided activities during their free-time. There were 295 completed questionnaires collected out of 409 handed out to visitors on numerous tour groups during this period, a return rate of 72%. Ten of these questionnaires were unusable due to incomplete data, leaving 285 completed surveys for the analysis. The cognitive and affective indicators were calculated by averaging the sum of the item scores (see Table 1) and are represented in a series of bar graphs that illustrate the number of responses in relation to a 'low' to 'high' impact, with the range of scores varying depending on whether there were three, four or five items associated with each indicator (again, see Table 1). SPSS data analysis software was used to construct each scale and to determine the alpha coefficients.

The methods and instrument used in this study do have some limitations. Firstly, the data are descriptive in nature and do not reveal reasons or causes for a particular indicator's score. Secondly the instrument relies on visitors' perceptions of the impacts of the interpretation. In other words, they are self-reported measures of outcomes rather than observed or evidence-

based measures of what visitors learned, felt or did as a result of the interpretation. Thirdly, the methods do not provide information about the longer-term impacts of interpretation on visitors once they have returned home. Fourthly, the heritage questionnaire developed as part of the *Interpretation Evaluation Tool Kit* is quite intentionally a generic instrument and the outcomes it assesses are not customized to the Sovereign Hill experience. Finally, this study is not a cross-cultural comparison and cannot assess interpretation-based outcomes due to cultural differences. These limitations are considered when interpreting the findings for the purposes of recommending changes to the tour and are revisited in the discussion of the findings of this study.

On the other hand, a strength of these methods is that, while the findings are population and site-specific, the outcomes of interpretation as reported by these visitors can be compared to other groups and even other sites, although such comparisons need to be done with caution for a few reasons. Firstly, the experience of the Chinese tour group visitor is very different to that of other Sovereign Hill visitors who typically spend all day at the site as a family or part of a school group visit; the Chinese visitor's experience is much shorter and more controlled compared to a free independent traveller's experience. Secondly, there is some suggestion that Chinese survey respondents are less likely to use the extreme (one and seven) responses categories in scale items, regardless of how extremely positive or negative they may feel due to cultural norms for central tendency when making evaluations and comparisons (Doosje et al., 1998). The next section presents the results of the instrument validation process, followed by the results and discussion of visitors' self-reported outcomes.

Results and Discussion

This section provides results of the data analysis. Part one (i) of these results discusses the reliability-testing of the translated instrument. Part two (ii) outlines the findings, beginning with the cognitive and affective indicators of interpretation presented as mean scores (on a scale of one to seven). Following the multi-item indicators, the findings associated with the behavioral questions (each of which was measured using a single item) are discussed and presented as a composite bar graph.

(i) *Results on scale reliability testing*

Previously developed scales may save researchers time and effort, however, they still need to be validated, as direct translations do not ensure an accurate representation of original content (Brislin, 1970; Cha, Kim, & Erlen, 2007). Though the English version of the instrument used in the present study has previously been established as a valid and reliable instrument, it was necessary to assess the reliability of the Chinese version due to the translation process. Cronbach's alpha coefficient scores were used to determine if the indicators in the translated version of the instrument demonstrated sound psychometric properties. The Cronbach's alpha coefficient is a determinant of the scale's internal reliability (how the items 'hang' together) on a scale of zero to one, with a larger value indicating greater internal reliability (Pallant, 2001). Each scale was constructed using SPSS data analysis software and the alpha coefficients were determined. All of the translated indicators demonstrated acceptable levels of internal reliability with the exception of the scale measuring the interpretation's relevance and meaningfulness. The *Positive attitude toward heritage* indicator Cronbach's alpha was less than .70 and though not ideal (Tabachnick & Fidell, 2007), it was still considered within

the limitations of acceptability for the purpose of this study. Each scale and their Cronbach's alpha coefficient scores are presented in Table 1.

(ii) Results on indicators measuring cognitive, affective and behavioral outcomes

One of the concerns in evaluating Chinese visitors using quantitative measures was their reported reluctance in using extreme (one and seven) response categories. An examination of the frequencies of each item indicated that the respondents were not hesitant in evaluating interpretation in relation to the extreme category of response. For example, on one of the *Impact via Empathy* items (relating to people who lived in the period being interpreted) 22.5% of respondents selected the lowest extreme response and 20% of respondents selected the highest extreme response.

Beginning with the outcomes that were mainly cognitive in nature, the results suggest that the impact was average to high. *Impact via Empathy* (Impact on current world view via empathy with historic period and people – Indicator A in Table 1) measures the impact that the interpretation had on visitors' views of their own lives by getting visitors to empathize with historic people and events. On a scale of one to seven, the mean response for the sample was 3.7 and, as illustrated in Figure 1, approximately 50% of responses were in the mid-range. This score indicates an 'average' outcome on this mainly cognitive indicator, but could be considered reflective of the nature of the tour. The tour is relatively brief and may be difficult for the guides to develop empathy in such a short period of time.

Insert Figure 1 here

Elaboration (Indicator B in Table 1) assesses the degree to which the visitor is provoked to thought in relation to the tour. When visitors elaborate on what they have learned, they make connections with what they already know. It has been suggested that elaboration is the most

important outcome of successful interpretation and indicates how much the interpretation provokes new thoughts about the topic of the interpretation (Ham & Weiler, 2006). On a scale of one to seven, the mean response for the sample was 5.2, with over 75% of the sample in the mid to high-impact range (see Figure 2). This score indicates a good outcome suggesting that the tour is relatively successful in engaging the visitor on a cognitive level.

Insert Figure 2 here

Turning to the indicators of affective outcomes, impact of the interpretation on all three indicators can be described as being in the very good range. The *Positive Attitude toward Heritage* indicator (Indicator C in Table 1) assesses the degree to which the interpretation leaves visitors with a sense of the importance of heritage preservation. On a scale of one to seven, the mean response on this affective indicator was 5.8, with approximately 50% of respondents in the high-impact range (see Figure 3). This suggests that the tour generally is very good at eliciting a positive attitude toward preserving heritage.

Insert Figure 3 here

The *Positive Global Evaluation* indicator (Indicator D in Table 1) assesses the fulfilment of the visitor's general expectations of enjoyment and satisfaction of the interpretation on the tour. On a scale of 1-7, the mean response on this affective indicator was 5.8. As illustrated in Figure 4, nearly 50% of respondents scored in the high-impact range, suggesting that the tour leaves a very good impression with visitors.

Insert Figure 4 here

The *Positive Word-of-Mouth* indicator (Indicator I in Table 1) assesses the willingness of visitors to promote the site based on their experience. On a scale of one to seven, the mean

response on this mainly affective indicator was a very high 5.7, with nearly 90% of responses in the mid to high-impact range and 43% in the high-impact range (see Figure 5). This indicates that those who took the Chinese-language tour generally had a positive experience which they want to share with others.

Insert Figure 5 here

No interpretable results were obtained regarding the relevance and meaningfulness indicator (Indicator J in Table 1) due to the fact that the scale did not reach the requirements for statistical reliability. There appears to be a need for further development and testing of items for this indicator in a Chinese-language context.

The remaining results are responses to yes/no questions in relation to the behavioral intentions of the visitors after completing the interpretive tour (see Indicators E, F, G and H in Table 1). Figure 6 generally reveals a modest but favorable impact on these four indicators, with 75% indicating a desire to stay longer, 69% desiring to purchase a souvenir related to the site story, 68% wishing to participate in additional interpretive activities, and 52% indicating a desire to return. The responses to the latter question may be partially due to the limited opportunity for overseas travel that most Chinese visitors expect to have.

Insert Figure 6 here

The summary statistics for all the indicators are presented in Table 2 from highest to lowest. While it is not possible to make direct comparisons between means and frequencies, the table suggests that affective outcomes were impacted most by the interpretation, and cognitive and behavioral outcomes were impacted less. As foreshadowed earlier, further examination of the interpretation itself would be needed to ascertain what might explain these outcomes. In addition, comparisons between these visitors to other Sovereign Hill visitors might provide

insight as to how the tour experience differs in its impact compared to visiting Sovereign Hill as an independent traveller. Notwithstanding the fact that comparisons of survey results across cultures can be fraught with difficulties, such comparisons are facilitated with a standardised instrument such as the one used in this study.

Insert Table 2 here

Conclusion and Implications

Although there is much hyperbole in the literature about the impacts of interpretation on visitors' beliefs, knowledge, understanding, attitudes and behaviors, the evidence to date is largely anecdotal or is site or activity specific. In other words, external validity and reliability in these studies has been limited. Tourism operators and managers need to be able to assess the extent of such impacts more rigorously, and to compare outcomes across visitor groups and across interpretive products. This study has provided sound evidence that the translated version of the heritage questionnaire provides practitioners and researchers with a valid and reliable tool (the *Relevant and Meaningful* indicator excluded) to evaluate interpretation-based outcomes in a Chinese language context.

Additionally, the results of this study suggest that the interpretation delivered as a part of the Chinese-language tour groups at Sovereign Hill is moderately successful in eliciting a range of identified cognitive, affective and behavioral outcomes. The visitor responses on the *Impact via Empathy* indicator suggest that the tour does what might be considered an 'average' job of impacting visitors' views of their own lives. This result may be reflective of the relative brevity of the tour, possibly not allowing the guide to develop strong themes during their interaction with visitors. The tour did, however, elicit 'good' visitor responses in relation to *Elaboration* suggesting that the tour is successful in engaging visitors on a cognitive level. A strong result was also found in relation to visitors' *Attitudes toward Heritage Preservation*

with around half of visitor responses in the high-impact range. Another strong result was found in relation to visitors' *Positive Global Evaluation* of the tour also having approximately half of the visitor responses in the high-impact range. The positive results generally associated with the tour were reflected in visitors' willingness to promote the site based on their experience; the majority of visitor responses on *Positive Word-of-Mouth* were in the mid to high-impact range. The behavioral questions included in the survey also elicited some positive results. The final indicator, *Relevant and Meaningful*, delivered no interpretable results due to a failure of the scale to reach accepted levels of statistical reliability.

The results of the present research need to be treated with a level of caution, as the instrument, as stated at the outset, is a generic one, not customised to the goals of Sovereign Hill and its interpretive program. Furthermore, the instrument is not based on evidence of Chinese visitors' pre-visit expectations of what outcomes they will gain from their visit. As mentioned earlier in this paper, a more complex research design is required to assess cause-and-effect relationships and to establish what specific actions need to be taken to improve interpretation. In other words, the present study does not provide answers as to *why* interpretation is or is not achieving a tourism attraction's desired outcomes. There is a growing body of evaluation research that investigates why an interpretive product (e.g. a guided walk, tour, interpretive talk, demonstration or other form of interpretive activity) is or is not effective or "successful" (Knapp, 1996; Knapp & Benton, 2005; Madin & Fenton, 2004). For example, in the case of face-to-face interpretation, success factors might include the message, the interpreter's personality, competence, personal performance, the physical environment or settings, the program content, and many other elements (Ham, 1986; Ham & Weiler, 2005; Yamada & Ham, 2004). Findings about these *input* variables can be useful for setting realistic objectives (Henderson & Bialeschki, 1995), designing interpretive experiences (Hayward & Brydon-Miller, 1984), allocating resources (Beckmann & Savage, 2003; Weinreich, 1999), sharing

what was learned with others (Jakobsen, 1999), and identifying what action needs to be taken for improvement (Belch & Belch, 2001). These are important areas of research that were beyond the scope of the present study but could be incorporated into the design of future research using this study's research instrument. Finally, the instrument was developed to assess interpretation in western contexts and reflect the values and perspectives and aims of western-based tourism. Though the instrument was, generally, successful in measuring Chinese-speaking inbound tourists' evaluation of western attempts at interpreting an Australian heritage tourism site, it does not capture whether the interpretation reflects eastern tourists' expectations, values and perspectives. Further research examining the cultural variations in expectations, values and perspectives is required.

In conclusion, the results of this study offer evidence, notwithstanding the fact that work still needs to continue in refining the instrument particularly in relation to the *Meaningful and Relevant* indicator, that the translated version of the heritage questionnaire provides practitioners and researchers with a reliable and valid tool to assess interpretive programs aimed at inbound visitors from China. In addition to this methodological contribution, the study is the first known published research that evaluates the impact of a Chinese-language interpretive tour. Results do, however, indicate that the instrument reliably assessed Chinese visitors' perceptions of their interpretive experience at Sovereign Hill as average to very good on cognitive, affective and behavioral constructs.

Reference list

Allard, M., Boucher, S., & Forest, L. (1994). The museum and the school. *McGill Journal of Education*, 29(2), 197-212.

Allen, S. (1997). Using scientific inquiry activities in exhibit explanations. *Science Education*, 81(6), 715-734.

Anderson, D., Lucas, K., & Ginns, I. (2000). Development of knowledge about electricity and magnetism during a visit to a science museum and related post-visit activities. *Science Education*, 84(5), 658-679.

Ash, D. (2003). Dialogic inquiry in life science conversations of family groups in a museum. *Journal of Research in Science Teaching*, 40(2), 138-162.

Ballantyne, R., & Uzzell, D. (1999). International trends in heritage and environmental interpretation: Future directions for Australian research and practice, *Journal of Interpretation Research*, 14(1), 59-75.

Beckmann, E., & Savage, G. (2003). *Evaluation of dingo education strategy and programs for Fraser Island*. Brisbane, QLD: Queensland Parks & Wildlife Service.

Beeho, A., & Prentice, R. (1995). Evaluating the experiences and benefits gained by tourists visiting a socio-industrial heritage museum: An application of ASEB Grid analysis to Blists Hill Open-air Museum, the Ironbridge Gorge Museum, United Kingdom. *Museum Management and Curatorship*, 14(3), 229-251.

Behling O., & Law K.S. (2000) *Translating questionnaires and other research instruments: Problems and solutions*. Thousand Oaks, CA: Sage Publications.

- Brislin R.W. (1970) Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology* 1, 185–216.
- Belch, G., & Belch, M. (2001). *Advertising and promotion: An integrated marketing communications perspective* (4th ed.). New York, NY: McGraw-Hill.
- Bicknell, S., & Mann, P. (1993). A picture of visitors for exhibition developers. In D. Thompson, A. Benefield, S. Bitgood, H. Shettel, & R. Williams (Eds.), *Visitor studies—theory, research and practice volume 5* (pp. 88-98). Jacksonville, AB: Visitor Studies Association.
- Bitgood, S., Patterson, D., & Benefield, A. (1988). Exhibit design and visitor behavior: empirical relationships. *Environment and Behavior*, 20(4), 474-491.
- Bitgood, S., & Patterson, D. (1993). The effects of gallery changes on visitor reading and object viewing time. *Environment and Behavior*, 25(6), 761-781.
- Blud, L. (1990). Social interaction and learning among family groups visiting a museum. *Museum Management and Curatorship*, 9, 43-51.
- Boisvert, D., & Slez, B. (1994). The relationship between visitor characteristics and learning-associated behaviors in a science museum discovery space. *Science Education*, 78(2), 137-148.
- Camp, B., Koran Jr., J., & Koran, M. (2000). Photographs as a research tool in visitor studies. *Journal of Interpretation Research*, 5(2), 47-58.
- Cartmill, R., & Day, L. (1997). Prevention of substance abuse: Can museums make a difference? *Curator*, 40(3), 197-210.
- Cha, E. S., Kim, K. H., & Erlen, J. A. (2007). Translation of scales in cross-cultural research: issues and techniques. *Journal of Advanced Nursing*, 58(4), 386-395.

China National Tourism Administration (CNTA) (2001). *"15 Year" Planning for China Tourism Development & 2015, 2020 Long Term Targets*. Beijing: China National Tourism Administration.

Chiozzi, G., & Andreotti, L. (2001). Behavior vs. time: understanding how visitors utilize the Milan Natural History Museum. *Curator*, 44(2), 153-165.

D'Agostino, J.V, Loomis, R.J., & Webb, B. (1992). Attitudes, beliefs, intended behaviors, and exhibit evaluation. In A. Benefield, S. Bitgood and H. Shettel (Eds.). *Visitor studies: Theory, research and practice volume 4* (pp. 92-101). Jacksonville, AL: Jacksonville State University Center for Social Design.

Department of Industry, Tourism and Resources (2005). *Media Release*. Canberra, ACT.

Department of Industry Tourism and Resources (DITR) (2007). *China Approved Destination Status (ADS) Scheme*. Department of Industry Tourism and Resources. Retrieved February 18, 2007, from

<http://www.industry.gov.au/content/itrinternet/cmscontent.cfm?objectID=8406AE50-65BF-4956-BA90AC2C508A80DC&CFID=14539222&CFTOKEN=33928357>

de Rojas, C., & Camarero, C. (2008). Visitors' experience, mood and satisfaction in a heritage context: Evidence from an interpretation center. *Tourism Management*, 29(3), 525-537.

Dierking, L., Adelman, L.M., Ogden, J., Lehnhardt, K., Miller, L. & Mellen, J. (2004). Using a behavior change model to document impacts of visits to Disney's Animal Kingdom: A study of investigating intended conservation action. *Curator*, 47(3), 322-43.

Doosje, B., Haslam, S. A., Spears, R., Oakes, P. J., & Koomen, W. (1998). The effect of comparative context on central tendency and variability judgements and the evaluation of group characteristics. *European Journal of Social Psychology*, 28, 173-184.

- Falk, J. (1997). Testing a museum exhibition design assumption: Effect of explicit labeling of exhibit clusters on visitor concept development. *Science Education*, 81(6), 697-687.
- Fernandez, G., & Benloch, M. (2000). Interactive exhibits: how visitors respond. *Museum International*, 52(4), 53-59.
- Finson, K., & Enochs, L. (1987). Student attitudes toward science-technology-society resulting from visitation to a science-technology museum. *Journal of Research in Science Teaching*, 24(7), 593-609.
- Ham, S. (1992). *Environmental interpretation: A practical guide for people with big ideas and small budgets*. Golden, CO: Fulcrum Publishing/North American Press.
- Ham, S., & Weiler, B. (2006) *Development of a research-based tool for evaluating interpretation*. Brisbane, QLD: Sustainable Tourism Cooperative Research Centre.
- Ham, S., & Weiler, B. (2007) Isolating the role of on-site interpretation in a satisfying experience. *Journal of Interpretation Research*, 12(2), 5-24.
- Harvey, M., Loomis, R., Bell, P., & Marino, M. (1998). The influence of museum exhibit design on immersion and psychological flow. *Environment and Behavior*, 30(5), 601-627.
- Hayward, G., & Brydon-Miller, M. (1984). Spatial and conceptual aspects of orientation: visitor experiences at an outdoor history museum. *Journal of Environmental Systems*, 13(4), 317-332.
- Henderson, K., & Bialeschki, D. (1995). *Evaluating leisure services: Making enlightened decisions*. State College, PA: Venture Publishing.
- Henriksen, E. (2001). High school students' understanding of radiation and the environment: Can museums play a role? *Science Education*, 85(2), 189-207.

Hernandez, S., & Celis Herrero, J. (1993). Beliefs in the science museum: An exhibit evaluation. . In A. Benefield, S. Bitgood and H. Shettel (Eds.). *Visitor studies: Theory, research and practice volume 6* (pp. 116-128). Jacksonville, AL: Jacksonville State University Center for Social Design.

Jakobsen, C. (1999). *Evaluation of interpretive exhibits in Rocky Mountain National Park*. Moscow, Idaho, USA: University of Idaho, Department of Conservation Social Sciences.

JuiChi, C. (2006). Customer satisfaction with tour leaders' performance: A study of Taiwan's package tours, *Asia Pacific Journal of Tourism Research*, 11(1), 97-116.

Knapp, D. (1996). Evaluating the impact of environmental interpretation: a review of three research studies. In *3rd Coalition for Education in the Outdoor Research Symposium Proceedings* (pp. 127-135). Bradford Woods, IN: Indiana University.

Knapp, D., & Benton, G. (2005). Long-term recollections of an environmental interpretive program. *Journal of Interpretation Research*, 10(1), 51-53.

Knudson, D., Cable, T., & Beck, L. (1995). Evaluation. In *Interpretation of cultural and natural resources* (pp. 441-465). State College, PA: Venture Publishing.

Korn, R.. & Jones, J. (2000). Visitor behaviour and experiences in the four permanent galleries at the Tech Museum of Innovation. *Curator*, 43(3), 261-281.

Leinhardt, G., & Crowley, K. (1998). *Conversational elaboration as a process and an outcome of museum learning*. Museum Learning Collaborative Technical Report (MLC-01). Pittsburgh, PA: University of Pittsburgh Learning Research and Development Center.

Leinhardt, G., Crowley, K., & Knutson, K. (Eds.). (2002). *Learning conversations in museums*. Mahwah, NJ: Lawrence Erlbaum Associates.

Leinhardt, G., & Knutson, K. (2004). *Listening in on museum conversations*. Walnut Creek, CA: AltaMira Press.

Light, D. (1995). Visitors' use of interpretive media at heritage sites. *Leisure Studies*, 14(2), 132-149.

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

McManus, P. (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(3), 263-270.

McManus, P. (1989). Oh, yes, they do: How museum visitors read labels and interact with exhibit texts. *Curator*, 32(3), 174-189.

Morgan, M., Lin, H., Chou, J. & Wu, H. (2006). An Interpretation Specialization Continuum of Environmental Volunteerism In Taiwan, *Journal of Interpretation Research*, 11(2), 7-20.

Pallant, J. (2001). *SPSS survival manual : a step by step guide to data analysis using SPSS for windows Version 10*, St Leonards, N.S.W.: Allen & Unwin.

Prentice, R., Sinead, G., & Stuart, M. (1998). Visitor learning at a heritage attraction: A case study of *Discovery* as a media product. *Tourism Management*, 19(1), 5-23.

ShiuhNan, H., Chuan, L., & HueiJu, C. (2005) The relationship among tourists' involvement, place attachment and interpretation satisfaction in Taiwan's national parks. *Tourism Management*, 26(2), 143-156

Stevens, R., & Hall, R. (1997). Seeing *Tornado*: How video traces mediate visitor

understandings of natural phenomena in a science museum. *Science Education*, 81(6), 735-747.

Stevenson, J. (1991). The long-term impact of interactive exhibits. *International Journal of Science Education*, 13(5), 521-531.

Stewart, Emma J., Hayward, Bronwyn M., Devlin, Patrick J., & Kirby, V. G. (1998). The 'place' of interpretation: a new approach to the evaluation of interpretation, *Tourism Management*, 19(3), 257-266.

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

Tourism Australia (TA) (2006). *China, Tourism Australia*. Retrieved January 21, 2007, from http://www.tourism.australia.com/content/MRRs/Dec06.v2/China_MMR_Dec06v2.pdf

Triandis H.C., & Brislin R.W. (1984). Cross-cultural psychology. *American Psychologist*, 39: 1006–1016.

Uzzell, D. (1998a). Strategic considerations and practical approaches to the evaluation of heritage interpretation. In D. Uzzell & R. Ballantyne (Eds.), *Contemporary issues in heritage and environmental interpretation: Problems and prospects* (pp. 185-202). London: Stationery Office.

Uzzell, D. (1998b). Evaluating the effectiveness of heritage and environmental interpretation. In Uzzell, D. & Ballantyne, R. (Eds.), *Contemporary issues in heritage and environmental interpretation: Problems and prospects* (pp. 203-231). London: Stationery Office.

Wearing, S. L., Archer, D., Moscardo, G., & Schweinsberg, S. (2006). *Development of a framework and casebook of best practice interpretation research that contributes to sustainable tourism*. Brisbane, QLD: Sustainable Tourism Cooperative Research Centre.

- Weiler, B., & Yu, X. (2007). Dimensions of cultural mediation in guiding Chinese tour groups: Implications for interpretation. *Tourism Recreation Research*, 32(3), 13-22.
- Weiler, B., & Yu, X. (2008). Case studies of the experiences of Chinese visitors to three tourist attractions in Victoria Australia, *Annals of Leisure Research*, 11(1&2), 225-241.
- Weinreich, N. (1999). *Hands-on social marketing: A step-by-step guide*. Thousand Oaks, CA: Sage Publications.
- White, T.G. and Jacobson, S.K. (1994). Evaluating conservation education programs at a South American zoo, *Journal of Environmental Education*, 25(4), 18-22.
- Yamada, N., & Ham, S. (2004). Assessing interpretation: Evaluative concepts and application to natural parks. *Japanese Journal of Outdoor Education*, 7(2), 1-12.
- Zeidler, A., & Surber, J. (1999). Understanding topic, structure, and importance of information in a visual and verbal display. *The Journal of Experimental Education*, 67(2), 114-134.

Table 1. Indicators, reliability scores and outcome constructs of original and Chinese version of the heritage interpretation evaluation instrument

Indicator	Indicator Title	Number of items (and questionnaire item numbers)	Internal reliability (Cronbach's alpha)	Chinese Version Internal reliability (Cronbach's alpha)	Type of outcome
A	Impact on current world view via empathy with historic period and people	5 items	.77	.74	mainly cognitive
B	Elaboration (provoked to thought)	5 items	.83	.71	cognitive
C	Positive attitude toward heritage / heritage preservation	3 items	.85	.66	affective
D	Positive global evaluation of interpretation at site	4 items	.84	.82	affective
E	Desire to participate in additional interpretive activities	1 item	N/A	N/A	mainly behavioral
F	Desire to purchase a memento or souvenir related directly to site story	1 item	N/A	N/A	mainly behavioral
G	Desire to stay longer	1 item	N/A	N/A	mainly behavioral
H	Desire to return for repeat visit	1 item	N/A	N/A	mainly behavioral
I	Positive word-of-mouth advertising	5 items	.89	.90	mainly affective
J	Visitors found it relevant and meaningful to their lives	4 items	.67	.35	cognitive

Adapted from Ham and Weiler (2006)

Table 2. Summary of interpretation outcomes in order from most to least favorable

Indicators measured using 7-point rating scale items	Type of outcome	Mean score
Positive attitude toward heritage / heritage preservation	affective	5.8
Positive global evaluation of interpretation at site	affective	5.8
Positive word-of-mouth advertising	affective	5.7
Elaboration (provoked to thought)	cognitive	5.2
Impact on current world view via empathy with historic period & people	cognitive	3.7
Visitors found it relevant and meaningful to their lives	cognitive	NA*
Indicators measured using yes/no responses		% “yes”
Desire to stay longer	behavioral	75%
Desire to purchase a memento or souvenir related directly to site story	behavioral	69%
Desire to participate in additional interpretive activities	behavioral	68%
Desire to return for repeat visit	behavioral	52%

*see explanation in part (i) of the results

Figure 1. Graphical representation of *Impact via Empathy* range of scores

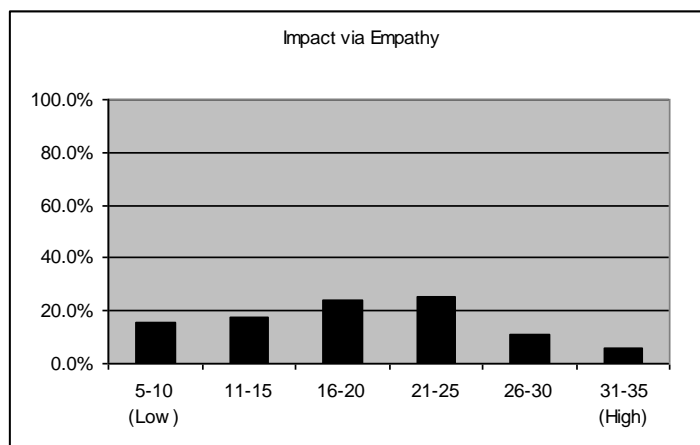


Figure 2. Graphical representation of range of scores for *Elaboration*

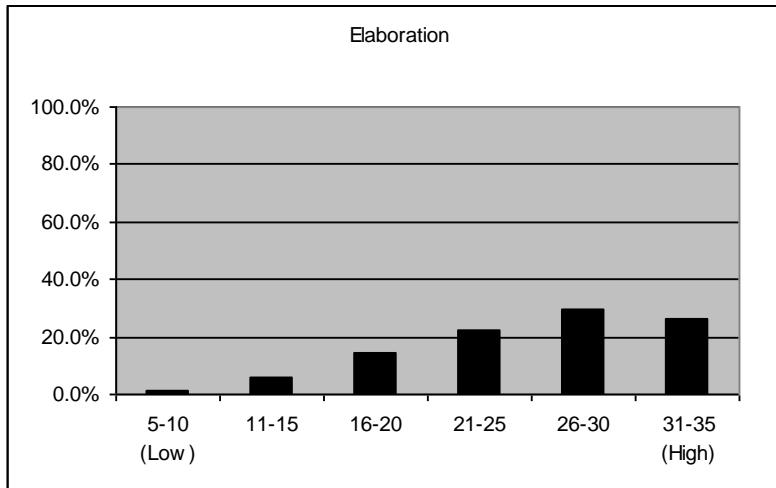


Figure 3. Graphical representation of *Positive Attitude toward Heritage* range of scores

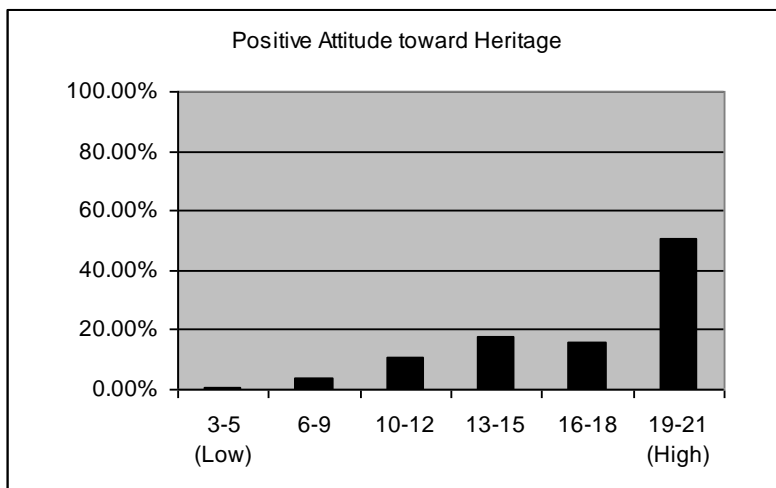


Figure 4. Graphical representation of *Positive Global Evaluation* range of scores

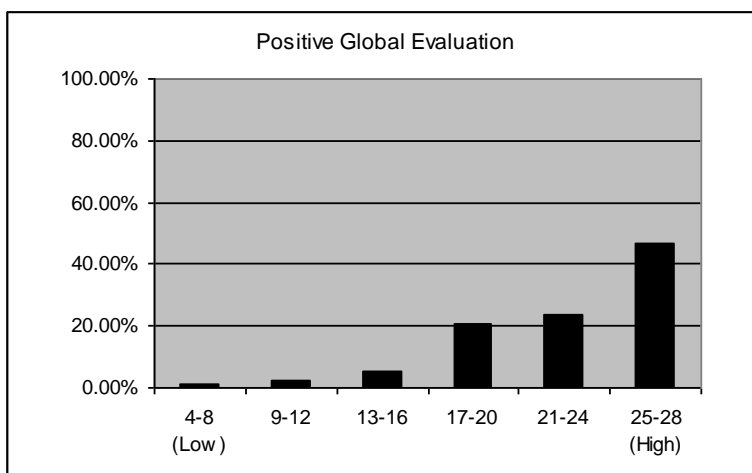


Figure 5. Graphical representation of *Positive Word-of-Mouth* range of scores

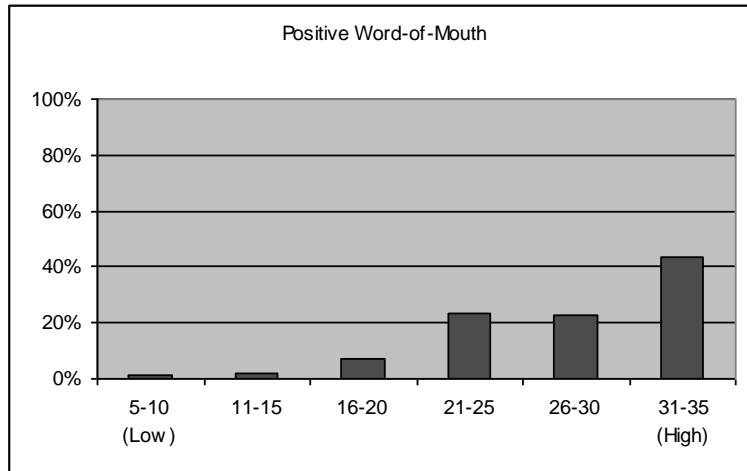


Figure 6. Graphical representation of favorable responses to *behavioral* outcomes

