The Effect of Audio Tours on Learning and Social Interaction: An Evaluation at Carlsbad Caverns National Park

LEVI T. NOVEY
115 S. Lilley Street, Apt. 101, Moscow, ID 83843, USA

TROY E. HALL
Department of Conservation Social Sciences, College of Natural Resources, University of Idaho, Moscow, ID 83844-1139, USA

Received 25 September 2005; revised 20 June 2006; accepted 24 July 2006

DOI 10.1002/sce.20184
Published online 12 October 2006 in Wiley InterScience (www.interscience.wiley.com).

ABSTRACT: Auditory forms of nonpersonal communication have rarely been evaluated in informal settings like parks and museums. This study evaluated the effect of an interpretive audio tour on visitor knowledge and social behavior at Carlsbad Caverns National Park. A cross-sectional pretest/posttest quasi-experimental design compared the responses of audio tour users ($n = 123$) and nonusers ($n = 131$) on several knowledge questions. Observations ($n = 700$) conducted at seven sites within the caverns documented sign reading, time spent listening to the audio, within group conversation, and other social behaviors for a different sample of visitors. Pretested tour users and nonusers did not differ in visitor characteristics, knowledge, or attitude variables, suggesting the two populations were similar. On a 12-item knowledge quiz, tour users' scores increased from 5.7 to 10.3, and nonusers' scores increased from 6.2 to 8.4. Most visitors were able to identify some of the park's major messages when presented with a multiple-choice question, but more audio users than nonusers identified resource preservation as a primary message in an open-ended question. Based on observations, audio tour users and nonusers did not differ substantially in their interactions with other members of their group or in their reading of interpretive signs in

Correspondence to: Levi T. Novey; e-mail: levi.novey@gmail.com; Troy E. Hall; e-mail: troyh@uidaho.edu

the cave. Audio tour users had positive reactions to the tour, and these reactions, coupled with the positive learning outcomes and negligible effects on social interaction, suggest that audio tours can be an effective communication medium in informal educational settings.
 Sci Ed **91:**260–277, 2007

INTRODUCTION

As technology continues to develop, it has become increasingly challenging for educators in informal settings to attract and hold people's attention, because people have become acclimated to dividing their attention among diverse varieties of competing stimuli (Sandifer, 2003). Moreover, in informal learning environments, people expect communications to get to the point quickly and are sometimes motivated by an interest in having fun (Falk, Moussouri, & Coulson, 1998). Although not universally the case, several studies have shown that many people are unwilling to devote sustained attention to media and messages that are not entertaining (Bitgood & Patterson, 1993; Falk, Phillips, & Boxer, 1993; Korn & Jones, 2000). In response, educators increasingly blend entertainment and learning through computers, video, and interactive activities (Gillies & Wilson, 1982; Sandifer, 2003). While novel technology can succeed in gaining people's attention, there is a need to determine whether these kinds of technology are effective at accomplishing their intended educational goals (Gramann, 2003).

In informal settings, such as parks, science centers, and museums, the challenge of education can be especially intense because the competition for attention is great and visitors are under no obligation to pay attention or learn. People often spend little time viewing interpretive and educational materials. Davidson, Heald, and Hein (1991) documented an average time of approximately 3 minutes spent in a natural history museum wing with six complex diorama exhibits; Sandifer (1997) noted that the typical museum visitor engaged with only 39% of the exhibits; and Bitgood, Pierce, Nichols, and Patterson (1987) reported that visitors spent an average of less than 200 seconds in a 1500-ft^2 simulated cavern containing a variety of formations and backlit signs. Additionally, many evaluations of the effectiveness of signs and exhibits confirm that people usually spend at most a few seconds reading labels or text, and learning can be quite variable. For instance, Chiozzi and Andreotti (2001) found that less than 7% of the exhibits in a museum of natural history were read comprehensively, and up to 70% received only cursory glances. In their cave study, Bitgood et al. noted that the average sign-reading time was less than 2 minutes, though approximately 25 minutes would have been necessary to read all of the material. In a study of learning from a visit to a historical museum, Prentice, Sinead, and Stuart (1998) found that between 34% and 82% of exiting visitors answered multiple-choice questions correctly. Such studies highlight the need for more effective nonpersonal forms of communication and to understand the factors that enhance learning outcomes.

One possibility for improving nonpersonal communication is the use of auditory modes. The theory of multiple intelligences proposes that everyone possesses some ability to learn through listening and music (Gardner, 1983, 1999; Nolen, 2003). Classroom studies have shown that aural components of teaching as well as visual methods are often effective for learners (Budoff & Quinlan, 1964; Harker & Feldt, 1993). However, while audio technology rapidly improves, it is still a rarity to hear music, sounds, or recorded narratives as interpretive tools at parks and visitor centers, although science centers are adapting more readily. Such media can be expensive to generate, and rigorous evaluations are needed to document their value in contrast to more traditional forms of communication.

The few studies of sounds and spoken text from informal science learning venues seem promising. For instance, auditory modes of communication have been shown to attract and stimulate visitor interest in a museum exhibit about birds (Peart, 1984), generate positive moods in a gorilla/jungle habitat in a zoo (Ogden, Lindburg, & Maple, 1993), and create interest in the history and features included in outdoor walking and driving tours (Wagar, 1976). Beer's (1987) study of a variety of museums found that exhibits with text and a model attracted the attention of 50% of visitors, while similar exhibits that included an audio component attracted the attention of nearly all visitors. Likewise, Davidson et al. (1991) observed that 72% of museum visitors listened to recorded labels, while only 34% read the accompanying written labels.

Although it appears that auditory elements in both sound-emitting exhibits and audio-recorded narratives attract attention, their effect on learning is less clear. For instance, Beer (1987) found that audio components did not sustain attention for more time than nonauditory exhibits. Peart (1984) reached a similar conclusion that sounds increased attention, but had no greater effect on knowledge than the exhibit without the audio. Such evaluative studies are rare; however, and there is a need for additional investigation of learning from audio communications. While it is well established that visitors rarely read all of the text on a sign, it is not known whether they will attend to an audio tour in its entirety or retain the information it contains. Our focus in this study is the audio tour used in Carlsbad Caverns National Park, and our first goal was to understand its effect on learning.

Evaluating Communication Effectiveness and Learning in Informal Settings

Evaluative studies use a variety of measures to judge success, and there is some debate about the most appropriate metrics. There are two interrelated issues: the definition of learning and the best methods to assess that learning. Some have used total time spent in exhibit spaces (e.g., Sandifer, 1997), which is fairly easy to document reliably and accurately. However, this is probably the most indirect and coarse way to draw inferences about cognitive change, because it tells little about visitors' actual engagement with materials. Other studies use the time spent reading labels or engaging in activities as an indicator of learning (Beer, 1987; Bitgood & Patterson, 1993; Korn & Jones, 2000). These typically describe both attracting power (the percentage of visitors who read or look at materials) and holding time (the average amount of time spent engaging with materials). Falk (1983) reported a correlation of 0.60 between time spent and knowledge gain in a science center, suggesting that time measures may be adequate proxies for learning in some instances.

Another approach is to measure cognitive change directly, though there are different points of view about what should be measured. Many evaluations have focused on the factual knowledge visitors retain at the end of their visit (e.g., Brown & Koran, 1998; Falk et al., 1993; Jacobson, 1988; Lee & Balchin, 1995). If the communicator's goal is to convey information, such measures are quite informative, as long as the question content adequately covers the important material. Certainly, a direct influence on beliefs is a prerequisite for more indirect effects on attitudes or behavior (Fishbein & Yzer, 2003). However, knowledge quiz questions do not necessarily indicate long-term retention of information or integration of new knowledge into existing cognitive structures.

Recently, many communication and education specialists have endorsed constructivist learning principles in a context of free-choice learning, claiming that visitors construct their own meanings from visits to educational settings, and factual knowledge tests do not sufficiently assess this type of learning (Brody, Tomkiewicz, & Graves, 2002; Rennie & Johnston, 2004; Tofield, Coll, Vyle, & Bolstad, 2003). Free-choice learning broadly reflects a learner-based approach toward how people process information rather than a transmission–reception model of learning (Lindauer, 2005). An individual's background and personal interest in understanding science and other subjects are viewed as the primary catalysts leading to learning, rather than a simple transmission and absorption of specific messages via media or personal contact. To understand how visitors make meanings in informal education settings, more open-ended or qualitative approaches are needed, and scholars have been experimenting with interviews, personal meaning mapping (Falk & Storksdieck, 2005), and other techniques (Ash, 2003). A downside of this type of assessment is that it can be relatively time consuming and intrusive for visitors.

Considering the large amount of attention to didactic approaches in the past, a shift in emphasis toward constructivist learning makes sense today, and indeed there is ongoing debate among communication professionals about what types of goals should be set (Lindauer, 2005). For this study, we melded didactic and constructivist perspectives, in part to enhance our ability to draw general conclusions (facilitated by a quantitative, didactic approach), but also to understand how people draw their own conclusions (constructivist approach). In addition, audience knowledge of our topic (caves) and the effectiveness of the communication device (an audio tour) are not well understood, and the U.S. National Park Service has established specific goals related to visitor knowledge about cave resources and their protection. An evaluation that combined traditional methods of assessing knowledge in the form of multiple-choice questions with a constructivist component in the form of an open-ended question to understand meaning seemed appropriate for the context.

Potential Drawbacks of Audio Tours

Even if audio tours enhance learning of the material they present, their effects on other learning-related behaviors are unknown, and this needs to be explored in order to weigh the overall merits of developing audio tours. It is possible, for example, that they influence whether visitors read informational signs. If visitors think that the audio tour conveys all the important information, they might not be inclined to read signs and could potentially miss important information. If this were the case, communication specialists would need to ensure redundancy of critical messages. On the other hand, the audio tour might pique visitors' interest in certain topics, and thereby increase reading behavior. In this case, the two media might be developed to complement each other.

Audio tours have a potentially important, though unproven, drawback for learning: they might discourage visitors from interacting with one another as people become involved in listening to the program or are confined by headphones (Brown, 2002; Serrell, 1996; Wagar, 1976). Social interaction provides one method by which visitors can enjoy and remember their educational experiences (Ash, 2003; McManus, 1987). It also may be important to how some people learn and process information (Packer & Ballantyne, 2005; Piscitelli & Weier, 2002). As shown in Brockmeyer, Bowman, and Mullins' (1983) study of sensory interpretation, many senior citizens want social interaction as part of their experience with interpretive communications. Likewise, people of all ages enjoy social interaction with other visitors, as well as with interpreters or staff. Brown (2002, p. 309) argues that "a major drawback of the audio tour is that it limits interactions between members of a visiting party, thereby stifling opportunities to share perspectives and experiences—even for those who are not taking the audio tour. It is an odd experience indeed to walk through a gallery full of people, eerily silent, interacting only with their headsets!"

The lack of prior investigation into the effects and effectiveness of audio tours leaves us unable to hypothesize whether the suspected learning benefits outweigh the potential detriment to social interaction and other learning behaviors. While previous research has looked at the relationships between interactive exhibits and social interaction, claims that audio tours limit social interaction have yet to be substantiated with research (Fernandez & Benlloch, 2000). Social interaction was therefore an important aspect of our research.

Study Purpose

Our study was designed to investigate (1) time spent in exhibit spaces and listening to the audio tour, (2) the effect of the audio tour on sign reading behavior, (3) knowledge gain caused by using an audio tour as opposed to traditional print and visual media, (4) the effect of audio tour use on within-group social interaction, and (5) audio tour users' attitudes about the tour.

METHODS

Study Area and Audio Tour

Located in southeast New Mexico, Carlsbad Caverns National Park provides the region with one of its most well-known and heavily visited tourist destinations, with more than 450,000 visits annually in recent years. Carlsbad has gained its reputation due to the size of its caves and their beautiful formations and as a home to a large colony of Mexican free-tailed bats. A paved path descends from the natural cave entrance to the main cave, called the Big Room, where it joins another paved loop accessed via an elevator. Together, the trails have approximately 40 backlit interpretive signs placed at varying intervals that interpret the human history, geologic history, and scientific interest of the cave and its formations.

The audio tour at Carlsbad Caverns uses a handheld wand that is placed against the ear like a cellular phone. Along the paved trails, each of the 50 audio stops is marked with a small numbered sign. A speaker in the wand transmits information, cued when the user presses the stop number on the keypad. The Big Room tour includes 30 minutes of total narration. Approximately 20% of park visitors rent the audio tour, which is available in English or Spanish for a price of $3.00. Some audio stops are at the same places where interpretive signs are located, while others are not. The content of the audio tour mostly matches what can be found on the signs within the caves, but offers greater detail and touches on several subjects not addressed in the signs. Sounds and music are used where appropriate to add to the program's quality. A female and a male actor alternate the narration in a conversational format. They occasionally ask questions of listeners to help create interest and foreshadow information that will be provided at upcoming audio locations. Interview excerpts with park staff, cave experts, and visitors add a diversity of voices to the program.

Evaluation Design

Many existing evaluations of informal educational materials utilize research designs with inherent weaknesses. For example, some (e.g., Prentice et al., 1998) do not include a pretest, so it is not possible to discern the extent of learning from preexisting knowledge. Others rely on self-reports of learning (Ogden et al., 1993; Packer & Ballantyne, 2005), which are often inaccurate (cf. Espiner, 1999). The Carlsbad evaluation had two components, together intended to overcome such limitations of past research: a self-administered questionnaire and unobtrusive observation of visitor behavior.

The questionnaire component followed a pretest/posttest quasi-experimental design (Creswell, 2003; Graziano & Raulin, 2004). Audio tour users were the experimental group and compared to a control group of nonusers. This study component was cross-sectional, with different samples of both groups surveyed immediately before they descended or immediately after they emerged from the caverns. Pretesting was necessary to determine whether the two populations differed on any characteristics that might influence the conclusions of the study and to establish baseline knowledge. For example, it seemed possible that audio tour users might be predisposed to like technology, or they might enter the study having already learned more about Carlsbad than nonusers. If the pretest showed that the groups did not differ in such ways, then any observed differences in the posttest could more confidently be attributed to the audio tour. Different samples of visitors were pre- and post-tested to ensure that there would be no pretest priming effect.

Questionnaires

The questionnaires were four pages long and took an average of 5–10 minutes to complete. To develop a profile of audio users and assess their commonality with nonusers, we included questions about visitors' use of the park (number of past visits and interpretive activities), demographic information (age and gender), and preferences for learning and technology. To investigate learning, 12 multiple-choice questions were asked, similar to the approach used by Jacobson (1988) and Prentice et al. (1998). Their content covered material contained in the interpretive signs and the audio tour and dealt with factual material such as park history or the geologic processes that formed the caverns. The correct answer to most of the questions was available from both media, but two questions addressed things only mentioned in the audio tour. "Don't know" responses were available as an option, and these were scored as incorrect. An open-ended question asked, "What do you think is the most important thing the National Park Service wants visitors to know about Carlsbad Caverns?" This was an important test of whether visitors grasped the overall message (or theme) of the park, which some argue is a more informative assessment of learning than a knowledge quiz (Davidson et al., 1991). Posttest questionnaires asked respondents what proportion of the signs they had read. Finally, audio users were asked 13 additional questions about their reactions to the audio tour and its effect on social interaction in their group. These were posed as 7-point agree/disagree rating scales. Audio users also indicated how much time they spent listening to the audio tour and what percentage of stops they heard.

Observational Measures

Although the tour users' posttest asked about listening, sign reading, and social interaction, we did not want to rely on self-reports alone, and we developed observational measures of several behaviors. The dim cave environment, with its distinct paved path that confines visitor travel, and placement of backlit signs adjacent to the trail made an ideal setting for unobtrusive observation. We followed guidance from Webb, Campbell, Schwartz, and Sechrest (2000) to observe behavior within discrete spaces. Both audio users and nonusers were observed.

Pilot observations of several spaces were made to evaluate their adequacy for collecting data, and seven were ultimately selected: three numbered stops with only audio narration; two with both signs and audio stops; and two that had only signs. The pilot observations, made by the primary researchers, helped refine rules for recording data and ultimately generated high levels of intercoder reliability. Each space was defined with appropriate and clear boundaries to permit consistency in observations of different people.

Several behaviors were observed. First, we timed how long the target visitor spent in the defined space (using a stopwatch). Sign reading behavior was a critical variable, and four categories of reading were defined: not reading (no evidence of looking at the sign), glancing (a brief look at the sign while passing), partial reading (orienting to the sign, stopping, and reading, but for less than the time necessary to completely read it), and complete reading (spending at least the average amount of time needed to read the sign—determined by reading the sign at a normal pace). We also timed how long tour users actually listened to their wand while in each space.

We were interested in social interaction within groups, and we developed several operational measures. First, we noted whether or not the target person was talking or listening to another person or people while within the defined space. Second, we observed whether the target person deliberately touched or was touched by another group member. This could include things such as holding hands, a pat on the back, or other physical interaction. Third, we noted whether the target person pointed to signal to another person to look at a formation, feature, or sign. Finally, we noted whether the target person posed for or took a photograph that included other people while in the observation space. (We reasoned that a posed photograph is likely to signify the making of a memory that involves other people.) The pointing and photography measures were exploratory in nature as additional methods of operationalizing social interaction. During two sampling sessions, a nonuniformed park staff member simultaneously observed the same people as the researcher in the cave to establish the interrater reliability of the measures (reported in the Results).

The primary researcher did his best to act as a normal visitor, pretending to be an audio user, and not let people realize that he was observing them. Caves, by nature, are dark places and the deep shadows in the caves helped conceal the observer's behaviors. To the best of our knowledge, none of the observations were influenced by the observer's presence.

Data Collection

The evaluation took place over 2 weeks, from May 30 to June 12, 2004. This period encompassed two weekends, including the Memorial Day holiday weekend. The visitor population over this time was 22,881, and the number of audio units rented during this period was 4,662 (20.4% of visitors).

Observations and questionnaire distribution took place at different times on the same days, with timing alternated to ensure that both forms of data collection would be spread as equally as possible across different times of the day (early morning, late morning, early afternoon, and late afternoon). To choose people to approach for the questionnaire, a systematic sampling strategy was utilized. At the start of a questionnaire period, the researcher approached the first individual who passed a specific point at the entrance or exit to the cave and asked if he or she would be willing to fill out a questionnaire. If the person refused, the next person was approached. If that person agreed, he or she was given the questionnaire to complete, and the observer would wait for two more people to pass and select the second. Subsequently, he would ask the third person who passed, and then return to asking the first person. Rather than a standard next-to-pass process of respondent selection (Veal, 1997), alternating individuals in this way ensured that we did not systematically bias the sample toward group leaders. People who were not speaking English in their groups and visitors who were obviously distraught or having problems within their own group were not approached. The number of these groups totaled about 10 over the course of the study.

We established a target of 100 observations in each of the seven locations. Observational data were collected at different times of day over 13 days. Working from one location to another, the researcher selected target individuals by using the simple next-to-pass technique, similar to the approach used by Davidson et al. (1991). That is, after an observation was complete (the target moved out of the observation space), the next adult who entered the space was chosen for observation. If people entered the space side by side, selection altered between the person on the right and the person on the left. This technique assured a random sample and meant that the number of audio users and nonusers observed approximated their actual proportions in the visitor population.

Data Analysis

Questionnaire and observational data were analyzed using SPSS. To compare users and nonusers, chi-square tests, *t* tests, and ANOVAs were performed as appropriate for the level of the data and the number of groups being compared. Following convention, an alpha level of 0.05 was adopted as indicating statistical significance. For ANOVAs, Scheffe tests were used to make post hoc comparisons among groups.

RESULTS

Sample Sizes, Response Rates, and Reliability of Observations

From all four surveyed groups (audio user pretest and posttest, nonuser pretest and posttest), 254 questionnaires were received, with 66% of visitors agreeing to participate. Response rates for the pretested groups were not significantly different between audio users (73%) and nonusers (63%), chi-square $(1, N = 186) = 2.37$, $p = .12$. However, exiting audio users participated at a significantly higher rate (80%) than exiting nonusers (54%), chi-square $(1, N = 203) = 13.86$, $p < .0005$. Possible reasons for this discrepancy are mentioned in the Discussion. People who declined to complete the survey sometimes offered reasons for their refusal. Among nonrespondents during pretest periods, 28% made a comment about being on a tight time schedule, and 15% of nonrespondents made this type of comment when declining to fill out the posttest questionnaire.

The total sample size for observations was 700 individuals, with 100 observations at each of the seven predetermined locations. Of all people observed, 164 were audio users, 469 were nonusers, and 67 were in groups where another group member had an audio tour. (For simplicity and clarity, these last users were not included in analysis, leaving an effective sample size of 633.) On two occasions, interrater reliability was established for 57 observed visitors with the aid of a second observer. Overall interrater reliability (agreement) across 14 measures for each subject was 91%. For sign reading behavior, interrater reliability was 91%, and for the four social interactions interrater reliability ranged from 84% to 100%.

Comparison of Tour Users and Nonusers

Pretested audio users and nonusers were compared on several variables, including gender, education level, group size and composition, past visits to caves, sources of information used in trip planning, knowledge about Carlsbad Caverns, interpretive programs and services used, and attitudes toward technology and learning. The only statistically significant difference related to respondents' intention to take a ranger-guided tour during their visits: 29% of the nonusers planned to take such a tour, compared to 13% of the audio users, χ^2 $(3, N = 125) = 4.656$, $p = .03$. These findings indicate that there were no major differences in the characteristics of the two groups at the pretest. Thus, differences that appear between audio users and nonusers at the posttest are likely to be due to different ways they use the cave and/or the audio tour.

TABLE 1
Mean Time Spent (Seconds) at Observed Locations With Audio Stops, Signs, or Both

Stop Title and Type	Audio Users	Nonusers	t	p
Speleothem audio (audio only)	94	45	−4.22	<.0005
Longfellow's bathtub (audio only)	60	28	−5.44	<.0005
Lechuguilla cave (audio only)	46	23	−6.85	<.0005
Crystal Spring Dome (sign + audio)	55	42	−1.80	.08
Lower Cave (sign + audio)	70	61	−0.80	.43
Draperies (sign only)	19	21	0.49	.62
Speleothem (sign only)	33	37	0.49	.63

Time Spent in Spaces and Listening to the Audio

Posttested audio users were asked how much time they had spent actually listening to the audio tour while in the cave. More than 75% of the respondents marked that they had listened for at least 1 hour. They were also asked to estimate the percentage of audio stops at which they listened to the tour. Nearly 78% of audio users claimed to have listened at 60% or more of the audio locations.

At the three audio-only spaces we observed, audio tour users spent more than twice as much time as nonusers (Table 1). Nonusers had little reason to stop in these spaces and generally were traveling through them. For the other spaces, both those with signs and audio or only signs, there were no significant differences in total time. In other words, using the audio tour did not appear to lower the time spent by audio tour users in nonaudio locations.

Audio tour users were observed listening to their wands at all locations, even those without audio stops (Table 2). In spaces with audio stops, more than 75% of people listened to at least part of the audio narration and the mean time was more than 30 seconds of listening. Although our samples of the different types of spaces are too small to draw firm comparisons, listening did not appear to vary across audio stops with and without signs.

TABLE 2
Time Spent Listening to Audio Narration in Observed Locations

Stop Title and Type	% Who Listened	Time Spent Listening (Seconds) Median	Mean	SD	Max.	Duration of Audio
Speleothem audio (audio only)	94.7	53.0	55.2	35.9	146	91
Longfellow's Bathtub (audio only)	76.9	29.0	33.0	26.1	80	73
Lechuguilla Cave (audio only)	95.5	27.0	33.4	20.8	70	81
Crystal Spring Dome (sign + audio)	80.8	31.5	39.1	34.1	132	86
Lower Cave (sign + audio)	83.9	41.0	44.7	33.8	140	39
Draperies (sign only)	22.7	0.0	3.5	6.9	21	
Speleothem (sign only)	84.4	12.5	15.6	16.4	56	

TABLE 3
Percent of Visitors Observed Reading or Not Reading Signs by Stop Type

	Sign + Audio[a]		Sign Only[b]	
Behavior	Audio ($n = 57$)	Nonuser ($n = 123$)	Audio ($n = 40$)	Nonuser ($n = 139$)
No reading	32	19	30	35
Glanced at sign	13	16	32	25
Partial reading	16	10	8	11
Read entire sign	39	55	30	30

[a]Chi-square = 6.00; $p = .11$.
[b]Chi-square = 1.15; $p = .77$.

Sign Reading

The four observed signs elicited a range of reading intensities by both audio users and nonusers (Table 3). The percentage who appeared to read the whole sign ranged from a low of 23% (nonusers, Draperies sign) to a high of 57% (nonusers, Crystal Dome sign). The percentage who ignored the signs entirely ranged from a low of 18% (nonusers, Lower Cave sign) to a high of 43% (audio users, Lower Cave sign). Although reading was highly variable across signs, observational data indicate that there were no statistically significant differences in the degree of sign reading by audio users and nonusers. For the stops with signs only, the percentage of visitors reading the entire sign was identical for both groups (30%). For the sign-and-audio stops, the percentage reading the entire sign was higher for nonusers (55%) than audio users (39%), but this difference was not statistically significant.

Effect of the Audio Tour on Visitor Knowledge

As evident in Table 4, both audio users and nonusers learned a substantial amount of park specific information during their time in the caverns. Although both groups scored approximately 50% correct on the 12-question quiz at the pretest, gains were especially notable among audio users. At the pretest, nonusers scored 6.2 correct answers and audio users scored 5.7, and the difference was not statistically significant, t (115) = 1.17, $p = .24$. At the posttest, nonusers' scores increased to 8.4, while users' scores increased to 10.3. Both gains were statistically significant, and the difference between audio users and nonusers was significant at the posttest, t (120) = 5.36, $p < .0005$. A two-way ANOVA showed a statistically significant interaction effect (model $F = 54.89$, $p < .0005$; interaction $F = 17.86$, $p < .0005$), demonstrating that the difference between pre- and post-tested audio users was greater than between pre- and post-tested nonusers.

It is important to note that the posttest difference is largely due to the inclusion of two questions whose answers were only provided on the audio tour. One dealt with the discovery and unique formations of Lechuguilla Cave, and the other dealt with research on the medicinal properties of a type of cave bacteria. For most other questions—whose answers were addressed in both print and audio–gains in knowledge were similar for the two groups. An additional analysis comparing scores without these two questions indicated that knowledge scores increased significantly for both groups from pretest to posttest, from approximately 5.5 to 8.7 for audio users and from 5.9 to 8.0 for nonusers (F (3, $N = 239$) = 35.38, $p < .0005$). Despite this similarity in factual learning, many more of the posttested audio users (84%) than nonusers (59%) recognized the overall message about

TABLE 4
Percentage of Respondents Answering Knowledge Questions Correctly

	Pretest			Posttest		
	Nonuser ($n=65$)	Audio ($n=60$)	χ^2	Nonuser ($n=66$)	Audio ($n=63$)	χ^2
Most common type of bat is Mexican free-tail	45.2	51.7	0.52	83.1	92.2	1.77
Bats navigate with echolocation	83.9	73.3	0.61	80.0	73.7	0.69
Jim White was the first explorer	41.9	25.0	3.92*	76.6	98.2	12.35***
Damage caused by touching formations	91.7	91.4	0.01	97.0	100.0	1.76
Stalactite vs. stalagmite	70.5	59.3	1.65	83.3	92.9	2.55
Geological processes that formed Carlsbad	51.7	54.2	0.08	72.3	83.9	2.34
Year Carlsbad was designated a National Park	26.7	18.6	1.09	72.4	70.2	0.25
Aware of Lechuguilla Cave	10.0	3.4	2.07	19.7	75.9	39.21***
Depth of caverns	65.0	50.8	2.45	86.4	86.2	0.00
Cave bacteria research	16.7	11.9	0.56	25.8	84.5	42.76***
Why coins in pools are harmful	56.7	67.8	1.56	88.3	87.9	0.53
Significance of Carlsbad	53.2	45.0	0.83	59.1	84.5	9.66**

*$p < .05$. **$p < .005$. ***$p < .0005$.

the importance of Carlsbad: it is significant for the beauty and unique formations within its large caves and is the home to a world-famous colony of bats.

Prior to the knowledge quiz, the questionnaire asked an open-ended question similar to the final multiple-choice knowledge question, in which visitors were asked to state in their own words the main intended park message. These were classified into several broad categories. Two independent coders classified all responses, and agreement was 89.6%. Posttested visitors were significantly more likely to supply an answer to the question than pretested visitors (Table 5), and those who gave an answer on the posttest scored significantly higher on the knowledge quiz ($M = 9.5$) than those who did not answer the question ($M = 8.4$), t (120) = 2.22, $p = .03$). Answers to this question were distributed across a few main themes, primarily the need to preserve the beauty and natural conditions of the park for future generations or the geological processes of cave formation. The two posttest groups did not differ in the percentage stating a main message related to cave beauty or bats, the two key National Park Service messages. However, a large difference between audio users and nonusers was found in the percentage stating a preservation/conservation theme, with audio users significantly more likely to describe such a message. Although only 21% of audio users stated such a theme at the pretest, this percentage rose to 58% at the posttest. It appears that the audio users, to a greater extent than nonusers, are obtaining information that leads them to a conclusion that the park is special and in need of preservation. This type of affective response may be a desirable outcome from the perspective of the National Park Service.

Effect of the Audio Tour on Social Interaction

Posttested audio users were asked, using a 7-point Likert-type scale, how strongly they agreed or disagreed with the following statement: "If I had not used the audio tour, I would

TABLE 5
Percent of Respondents Identifying Primary Park Message in Open-Ended Question

	Pretest		Posttest		
	Nonuser ($n = 65$)	Audio ($n = 60$)	Nonuser ($n = 66$)	Audio ($n = 63$)	p^a
Gave an answer	67.7	58.3	84.8	77.8	.005
Primary message identified:					
Preservation/conservation	44.7	21.1	37.5	58.0	.005
Formation/creation of caverns	25.5	28.9	33.9	14.0	.12
Beauty/wonder	23.4	13.2	21.4	20.0	.68
History of caverns or exploration	12.8	18.4	12.5	2.0	.09
Do not harm or touch	10.6	2.6	8.9	18.0	.13
Bats	8.5	13.2	7.1	0.0	.10
Other	10.6	18.4	33.9	20.0	.03

[a]Chi-square test.

have talked more to other people while in the Caverns.” The mean rating was 0.14 (close to neutral) with a large standard deviation of 2.03. Inspection of the frequency distribution for this question showed a pronounced bimodal distribution: 27.6% strongly disagreed (marking −2 or −3 on the rating scale), while 30.1% strongly agreed (+2 or +3 on the rating scale).

The observational data revealed few significant differences between audio users and nonusers for the social interaction measures (Table 6). There were no differences between groups for any measure at the sign-and-audio stops. However, at the sign-only stops, nonusers were significantly more likely than audio users to be touching other group members. Additionally, at the audio-only stops, nonusers were significantly more likely to be conversing.

Audio Tour Users’ Attitudes Toward the Carlsbad Caverns Audio Tour

The questionnaire asked audio users for their agreement (on a scale from −3 to +3) with several statements relating to the audio tour (Table 7). These reveal that the audio tour is

TABLE 6
Percent of Visitors Engaging in Social Interaction in Different Types of Spaces

	Sign + Audio			Sign Only			Audio Only		
Behavior	A^a ($n = 57$)	N^a ($n = 134$)	p^b	A^a ($n = 40$)	N^a ($n = 139$)	p^b	A^a ($n = 201$)	N^a ($n = 67$)	p^b
Conversation	81	81	.92	63	63	.99	61	75	.025
Photography	2	2	.77	0	2	.35	6	9	.47
Pointing	18	18	.96	8	20	.06	9	12	.48
Touching	12	15	.57	0	18	.004	6	11	.22

[a]A = Audio tour user; N = nonuser.
[b]Chi-square test.

TABLE 7
Posttested Audio Users' Agreement With Statements About the Audio Tour

	Mean	*SD*	% Agree	% Disagree
The Carlsbad audio tour was informative	2.63	0.70	98.2	1.8
I had fun using the Carlsbad audio tour	2.13	1.06	88.7	1.6
I did not mind carrying the audio tour equipment around	2.07	1.39	86.9	8.2
The rental price for the audio tour was reasonable	1.81	1.48	82.3	6.5
The music and sounds included in the audio tour program enhanced its quality	1.71	1.27	79.0	3.2
I was provided with sufficient training as to how to operate my audio tour equipment	1.07	1.87	58.2	16.4
The audio tour provoked my interest more than signs would have	1.07	1.89	62.5	17.9
When using the audio tour I read signs less than I normally would	−1.10	2.07	27.4	64.5
I feel like I did not learn from the audio tour	−2.23	1.25	3.8	92.3
The audio tour equipment was difficult to use	−2.26	1.24	4.9	90.2
Using the audio tour made me feel tired	−2.45	0.99	86.2	3.4

well liked and perceived as easy to use. Audio users found it to be informative, and most (64.5%) disagreed that the audio caused them to read fewer signs. Two thirds (62.5%) said it provoked their interest more than signs would have done.

DISCUSSION

Audio tours are increasingly practical and user-friendly, as both hardware and software improve. Today there are several audio tour companies that help clients create professional interpretive programs that incorporate sounds, music, and narratives. Our findings suggest for several reasons that such programs are worth pursuing in informal educational settings like parks, zoos, science centers, and museums.

First, visitors enjoy the tour. Nearly all tour users found the audio to be informative and fun to use. Because there were no important differences between tour users and nonusers at the pretest, the reaction of tour users does not appear to be a function of self-selection bias. Instead, it seems likely that most users would find the tour enjoyable and engaging if they had the opportunity to try it.

Second, tour users listen extensively to the audio narration. Across all seven spaces we observed, audio users spent more total time. This was a result of their spending more time in the audio-only spaces and an equivalent amount of time as nonusers at the stops that had signs. In principle, the more time spent, the greater the potential for learning (Falk, 1983; Sandifer, 1997). If the audio increases the length of visits, as it seems to do in this case, managers of parks and other sites may judge this to be a beneficial outcome.

Audio users listened to narration in all spaces, even those without audio stops, and listening remained high at audio stops spaced from the beginning to the end of the Big Room trail. Thus, we did not observe the type of attentional fatigue sometimes seen in response to exhibits or signs in museums or galleries.

Third, audio tour users learn at least as much as visitors who do not use the tour. Other studies have documented a positive impact of sounds and audio on attention and mood (e.g., Ogden et al., 1993; Peart, 1984). However, the effects on learning have been less clear, and some research suggests that there is little advantage to audio (Beer, 1987). We did not find the audio tour to be superior in conveying information, at least for the topics included in our knowledge quiz, except for the obvious cases where information was only presented aurally. Nevertheless, the two items covered only in the audio tour were mentioned only briefly, yet approximately 80% of tour users answered those questions correctly at the posttest. Although the audio in and of itself did not convey an advantage in recalling facts, audio users were more likely to recognize the park's main themes in the multiple-choice question and to articulate a primary message of preservation in the open-ended question on the questionnaire. Therefore, the audio tour may have helped visitors integrate new understandings into an overall meaning.

Most evaluations of learning have been done at science centers or museums, with few from national parks or other protected area visitor centers. Therefore, comparisons of learning could be misleading, if people to different types of sites have different goals related to learning (Falk & Dierking, 2000). It seems plausible that people who go to science centers and museums may hope to learn, whereas people who visit parks may be focused on other experiences. To put our results in proper context, we reviewed studies conducted in similar types of settings. The magnitude of learning at Carlsbad Caverns—among both audio users and nonusers—seems substantial in comparison to results presented in these other studies. For instance, Jacobson (1988) found that correct scores on a 12-question test only increased from 3.9 to 6.3 among those exposed to signs and 6.6 among those who used a booklet during a visit to a nature trail. Brown and Koran (1998) studied learning among visitors to a Mayan ruin. Visitors' average scores on a 15-point multiple-choice test increased only minimally, from 8.7 to 10.2. In a science center, Falk et al. (1993) found that, on average 65% of exiting visitors could answer questions about six exhibits correctly, up from 28% of entering visitors. In our study, entering visitors scored approximately 50% correct, and exiting audio users averaged 86% correct. Although visitors' scores on knowledge quizzes obviously depend in part on the difficulty of the questions, the learning occurring at Carlsbad Caverns is on par with what is found in well-designed science centers.

Fourth, the audio tour does not appear to interfere with sign reading. Both tour users and nonusers devoted considerable attention to the signs in the cave, more than is often documented in studies of sign reading. This may be due to the way their colors stand out with backlighting in the generally dim and monochromatic cave environment. Additionally, cave formations are probably foreign to most visitors, who are therefore drawn to signs to help them learn about and interpret what they are seeing.

Finally audio tours—at least the type of handheld wand used at Carlsbad Caverns—do not appear to impede social interaction to a substantial degree. It appears that enjoyment and learning inspired by social interaction were not negatively influenced at Carlsbad Caverns by the audio tour. The two exceptions—which were not substantial—were for deliberate touching behavior at the sign-only stops and conversation rates at the audio-only stops. Probably the first difference is due to audio users having an audio wand in one hand and therefore being less able to touch other members of their group. The second difference is also unsurprising.

The need for social interaction to foster learning has been a subject of some debate in the literature, and the degree to which visitors desire social interaction in educational settings probably depends on the circumstances and individual differences. While social interaction is certainly important to learning outcomes in some settings, like science centers where parents interact with their children to develop understandings of complex concepts, it may be less important to learning in other settings. In a recent study, Packer and Ballantyne (2005) observed the time visitors to the Discover Queensland exhibit in Brisbane, Australia, spent looking at and reading materials. On average, people spent only 20% of their time interacting with others in their groups. Four weeks later people were queried about the number and type of memories they had, and there were no differences between people who had visited alone compared to those who had visited with others. The authors concluded that sharing experiences with others conveyed no significant learning advantage. In our study, observation showed that conversation and physical contact were affected to some degree, but factual learning was not impaired by the audio tour. Hence, we conclude that fear of adverse effects on social interaction may not be warranted.

Advantages of Audio Tours

While our study did not find a substantial improvement in learning from the audio tour beyond what was achieved by signs and other nonpersonal media, in suitable settings the audio tour is likely to be especially effective for several reasons. First, audio tours can be designed with flexibility, so that participants can follow paths that they choose (Brown, 2002; Hein, 1998; Serrell, 1996). With the touch of a button, users can also select the language in which they would like to hear a tour. This factor alone allows in-depth communication to reach visitors from different cultures. Tours can be produced for different age groups, and participants can choose the amount of detail they receive. Audio tours offer a medium for communicating information to people with certain disabilities (Anonymous, 1999a, 1999b). If, as seems to be the case, different people prefer different modes of instruction, inclusion of audio components in the set of communication tools can be advantageous. For instance, Davidson et al. (1991) showed that adding auditory, touch, and smell elements increased the holding time of a natural history museum wing from 3.1 to 5.3 minutes. If professionally produced and consistent with the best communication principles, audio tours should be more effective in reaching the broad public than other forms of nonpersonal communication.

Study Limitations and Future Research

We had uneven response rates among study groups and a significantly lower response by exiting people who did not use the audio tour. Although we can only speculate about reasons for this tendency, one possibility is that audio tour users—who listen to a narrative not included in the signs about the importance of research in the park—became more likely to want to be part of a research effort. However, it is also possible that the different response rates have to do with the way participants were approached. Audio users were told that the research was about learning from the audio tour, whereas nonusers were simply told that the research was to find out more about how much people learned while in the cave. Perhaps this caused audio users to feel special and want to contribute.

The lower response rate for posttested nonusers (54%) could introduce some bias in our conclusions. If people in a hurry or less captivated by the cave were systematically underrepresented, this could artificially inflate estimates of learning. However, there were no statistically significant differences between pretested and posttested nonusers on group characteristics, trip variables, or other factors, so we are unaware of any systematic bias in our findings.

Another limitation of our study was that we treated the questionnaire and observation components as largely independent, in that we surveyed and observed different people. In future studies, being able to link observations of individuals' behavior to their responses to knowledge or attitude questions would be valuable (e.g., Falk, 1983; Packer & Ballantyne, 2005).

Finally, our measures of social interaction were limited to those that we could observe with high reliability and carry out unobtrusively in the confined cave spaces. However, other researchers have demonstrated the value of more subtle measures of social interaction, such as documenting actual visitor conversations (McManus, 1989).

CONCLUSION

Ecologically relevant sounds have been shown to induce positive emotional reactions in visitors (Ogden et al., 1993). This study provides evidence that auditory communications that combine sounds and an audio-recorded narrative can also have a positive effect on cognitions. Most visitors who used the audio tour gained substantial knowledge about the park and were more likely to understand the park's interpretive themes–primarily that Carlsbad Caverns is of worldwide fame because of its large caves that contain beautiful formations, its large bat colony, and its importance to scientific research. Audio users liked the content of the tour, and behaviors of interest to park staff—such as reading interpretive signs—did not appear to be impeded. While interpretive specialists have raised concerns that social interaction might be adversely affected by audio tours, these effects seem minimal at Carlsbad Caverns. We hope that this encouraging finding will prompt more parks and visitor centers to consider developing audio tours to communicate the cultural and scientific value of the resources they steward.

The authors wish to thank Carlsbad Caverns National Park and the Carlsbad Caverns Guadalupe Mountains Association for their support and permission to conduct this research, as well as Jeff Kimberling and Nick Sanyal for their academic support.

REFERENCES

Anonymous. (1999a). Consumers. (Pilot program of audio tours). Journal of Visual Impairment & Blindness, 93(4), 260–261.

Anonymous. (1999b). Resources. (New York Hall of Science audio tours). Journal of Visual Impairment & Blindness, 93(9), 611–612.

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

Beer, V. (1987). Great expectations: Do museums know what visitors are doing? Curator, 30(3), 206–215.

Bitgood, S., & Patterson, D. (1993). The effect of gallery changes on visitor reading and object viewing time. Environment and Behavior, 25(6), 761–781.

Bitgood, S., Pierce, M., Nichols, M., & Patterson, D. (1987). Formative evaluation of a cave exhibit. Curator, 30(1), 31–39.

Brockmeyer, F. M., Bowman, M. L., & Mullins, G. (1983). Sensory versus nonsensory interpretation: A study of senior citizens' preferences. Journal of Environmental Education, 14, 3–7.

Brody, M., Tomkiewicz, W., & Graves, J. (2002). Park visitors' understanding, values, and beliefs related to their experience at Midway Geyser Basin, Yellowstone National Park, USA. International Journal of Science Education, 24(11), 1119–1141.

Brown, K. (2002). Educational and other public programmes for exhibitions. In B. Lord & G. D. Lord (Eds.), The manual of museum exhibitions (pp. 297–315). Walnut Creek, CA: Altamira Press.

Brown, F. S., & Koran, J. J. (1998). Learning from ruins: A visitor study at Uxmal. Curator, 41(2), 121–131.
Budoff, M., & Quinlan, D. (1964). Auditory and visual learning in primary grade children. Child Development, 35, 583–586.
Chiozzi, G., & Andreotti, L. (2001). Behavior vs. time: Understanding how visitors utilize the Milan Natural History Museum. Curator, 44(2), 153–165.
Creswell, J. W. (2003). Research design: Qualitative, quantitative, and mixed methods approaches (2nd ed.). Thousand Oaks, CA: Sage.
Davidson, B., Heald, C. L., & Hein, G. E. (1991). Increased exhibit accessibility through multisensory interaction. Curator, 34(4), 273–290.
Espiner, S. R. (1999). The use and effect of hazard warning signs: Managing visitor safety at Franz Josef and Fox Glaciers (Science for Conservation Report 108). Wellington, New Zealand: Department of Conservation.
Falk, J. (1983). Time and behavior as predictors of learning. Science Education, 67(2), 267–276.
Falk, J. H., & Dierking, L. (2000). Learning from museums: Visitor experiences and the making of meaning. Walnut Creek, CA: AltaMira Press.
Falk, J., Moussouri, T., & Coulson, D. (1998). The effect of visitors' agendas on museum learning. Curator, 41(2), 107–120.
Falk, J., Phillips, K. E., & Boxer, J. J. (1993). *Invisible Forces* exhibition: Using evaluation to improve an exhibition. In D. Thompson, A. Benefield, S. Bitgood, H. Shettel, & R. Williams (Eds.), Visitor studies: Theory, research, and practice (Vol. 5, pp. 212–226). Jacksonville, AL: Visitor Studies Association.
Falk, J., & Storkskieck, M. (2005). Using the contextual model of learning to understand visitor learning from a science exhibition. Science Education, 89(5): 744–778.
Fernandez, G., & Benlloch, M. (2000). Interactive exhibits: How visitors respond. Museum International, 52(4), 53–59.
Fishbein, M., & Yzer, M. C. (2003). Using theory to design effective health behavior interventions. Communication Theory, 13(2), 164–183.
Gardner, H. (1983). Frames of mind: The theory of multiple intelligences. New York: Basic Books.
Gardner, H. (1999). Intelligence reframed: Multiple intelligences for the 21st century. New York: Basic Books.
Gillies, P., & Wilson, A. (1982). Participatory exhibits: Is fun educational? Museums Journal, 82, 131–135.
Gramann, J. (2003). Trends in demographic and information technology affecting visitor center use: Focus group report. Washington, DC: USDI National Park Service, Social Science Program.
Graziano, A. M., & Raulin, M. L. (2004). Research methods: A process of inquiry. Boston: Allyn and Bacon.
Harker, J. K., & Feldt, L. S. (1993). A comparison of achievement test performance of nondisabled students under silent reading and reading plus listening modes of administration. Applied Measurement in Education, 6(4), 307–320.
Hein, G. E. (1998). Learning in the museum. London: Routledge.
Jacobson, S. K. (1988). Media effectiveness in a Malaysian Park system. Journal of Environmental Education, 19(4), 22–27.
Korn, R., & Jones, J. (2000). Visitor behavior and experiences in the four permanent galleries at the Tech Museum of Innovation. Curator, 43(3), 261–281.
Lee, T., & Balchin, N. (1995). Learning and attitude change at British Nuclear Fuel's Sellafield Visitors Centre. Journal of Environmental Psychology, 15(4), 283–298.
Lindauer, M. (2005). From salad bars to vivid stories: Four game plans for developing "educationally successful" exhibitions. Museum Management and Curatorship, 20, 41–55.
McManus, P. M. (1987). It's the company you keep... the social determination of learning-related behavior in a science museum. International Journal of Museum Management and Curatorship, 6(3), 263–270.
McManus, P. M. (1989). Oh yes they do: How museum visitors read labels and interact with exhibit texts. Curator, 32(3), 174–189.
Nolen, J. L. (2003). Multiple intelligences in the classroom. Education, 124(1), 115–120.
Ogden, J. J., Lindburg, D. G., & Maple, T. L. (1993). The effects of ecologically relevant sounds on zoo visitors. Curator, 36(2), 147–156.
Packer, J., & Ballantyne, R. (2005). Solitary vs. shared: Exploring the social dimension of museum learning. Curator, 48(2), 177–192.
Picscitelli, B., & Weier, K. (2002). Learning with, through, and about art: The role of social interactions. In S. G. Paris (Ed.), Perspectives on object-centered learning in museums (pp. 121–151). Mahwah, NJ: Lawrence Erlbaum.
Peart, B. (1984). Impact of exhibit type on knowledge gain, attitude change and behavior. Curator, 27(2), 220–237.
Prentice, R. C., 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.

Rennie, L. J., & Johnston, D. J. (2004). The nature of learning and its implications for research on learning from museums. Science Education, 88, S4–S16.

Sandifer, C. (1997). Time-based behaviors at an interactive science museum: Exploring the differences between weekday/weekend and family/nonfamily visitors. Science Education, 81, 689–701.

Sandifer, C. (2003). Technological novelty and open-endedness: Two characteristics of interactive exhibits that contribute to the holding of visitor attention in a science museum. Journal of Research in Science Teaching, 40(2), 121–137.

Serrell, B. (1996). Exhibit labels: An interpretive approach. Walnut Creek, CA: AltaMira Press.

Tofield, S., Coll, R.K., Vyle, B., & Bolstad, R.. (2003). Zoos as a source of free choice learning. Research in Science and Technological Education 21(1), 67–99.

Veal, A. J. (1997). Research methods in leisure and tourism. London: Pitman.

Wagar, J. A. (1976). Cassette tapes for interpretation. Portland, OR: USDA Forest Service, Pacific Northwest Forest and Range Experiment Station.

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (2000). Unobtrusive measures (revised ed.). Thousand Oaks, CA: Sage.