Electronic Guidebooks and Visitor Attention

Allison Woodruff[(*)], *Paul M. Aoki*[(*)], *Amy Hurst*[(*)],[1] *and Margaret H. Szymanski*[(*)]

[(*)] Xerox Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, CA 94304 USA

ABSTRACT
We describe an electronic guidebook prototype and report on a study of its use in a historic house. Supported by mechanisms in the guidebook, visitors constructed experiences that had a high degree of interaction with three entities: the guidebook, their companions, and the house and its contents. In this paper, we report a qualitative analysis of how different properties of the guidebook helped or hindered visitors' attempts to balance the competing demands of these attentional entities. Based on the visitors' comments and behavior, we distill a set of design principles.

KEYWORDS: Electronic guidebooks, visitor studies, user studies, evaluation

INTRODUCTION

Visitors to cultural heritage locations have multiple goals. Often they want to get information about the objects they see. However, sharing the experience with their companions is often a higher priority than education, particularly for infrequent visitors [10]. Similarly, a "romantic" experience, one in which the visitor "gets the sense of the place," is often prioritized more highly than education by both curators and visitors [2]. Therefore, at least three entities demand the visitor's attention at a cultural heritage site: (1) an information source (e.g., a guidebook); (2) their companions; and (3) the location itself. Moreover, each entity presents different stimuli that vary in attractiveness at different times. For example, a companion offers conversation, sometimes making observations of high interest to the visitor, sometimes not. As other examples, some objects are of higher interest to the visitor than others, and specific guidebook content may or may not discuss topics that the visitor wants to know about.

To manage the constantly changing demands of these entities, visitors perform a sophisticated balancing act. At any given moment, visitors divide

[1] Work performed during an internship from the College of Computing, Georgia Institute of Technology.

their attention in accordance with their priorities. They may attend to multiple things at once, or they may give their undivided attention to some particularly interesting stimulus. When a stimulus ceases or becomes less attractive, they may switch their attention from one entity to another, e.g., if a companion starts to say something boring, they may switch their attention to an object in the room. The ability to assess the attraction of different stimuli and switch contexts allows visitors to create engaging, fulfilling experiences.

Guidebooks and tours have the potential to help or hinder visitors as they strive for optimal attentional balance. Some guidebooks are known to have helpful qualities, e.g., ones that allow the visitor to choose which objects they learn about. Other presentation methods are known to hinder the visitors' attempts to balance their attention optimally. For example, a fully scripted audio tour may constantly demand the visitors' attention, never giving them the opportunity to look at objects other than the ones specifically mentioned in the tour or to converse with their companions. Such a tour denies the visitor the chance to balance and rebalance their attention. Visitors have different responses to this situation, depending on their priorities. Some visitors simply refuse to use such audio tours because they do not provide enough control, or because they isolate visitors from their companions. Other visitors prize the information or experience (e.g., a sense of continuity) offered by such audio tours highly enough to sacrifice personal choice and social interaction, but subsequently express a sense of discontent with their overall experience.

As designers of electronic guidebooks, we are interested in identifying guidebook properties that help visitors reach the desired attentional balance. To this end, we conducted a qualitative study of visitors. We first constructed a prototype, designing it to provide a range of options for information presentation and sharing. We then observed fourteen visitors using the guidebook in a self-guided tour of a historic house and conducted semi-structured interviews.

In this paper, we report a qualitative analysis of how different properties of the device helped or hindered visitors' attempts to balance the competing demands of multiple attentional entities. Based on the visitors' comments and behavior (e.g., how visitors used some features and ignored others), we distill a set of design principles. The analysis presented here focuses on content from the interviews, supplemented by the informal observations of visitor use of the guidebooks.

As might be expected, some of our findings were consistent with prior quantitative and qualitative work in museum studies. However, because our design afforded novel possibilities for interaction between the visitors, we were able to make novel observations about the critical effect that the guidebook had on visitors' sharing of their experience. Furthermore, careful design allowed us to provide a visual user interface that was robust and easy to explain yet lightweight. Ease of use significantly changed visitors' experiences. Finally, our model of attention management gave us interesting perspectives on "known" results.

In the next section, we describe our guidebook prototype. We then describe our research methods. We next describe the visitors' behavior with the guidebooks and then discuss their comments on their experience. After synthesizing design principles, we

discuss related work and conclude.

PROTOTYPE

The electronic guidebook application runs on a Casio Cassiopeia™ E-105 personal digital assistant (PDA), a small device weighing 255g (9 oz.). Its display is a color touch-sensitive screen. A user generally holds the PDA in one hand and a stylus in the other hand, touching the stylus to the screen to interact with the device.

Visitors obtain information about objects in their environment using a visual interface. The interface is akin to a set of Web browser imagemaps but has many refinements [3] that simplify operation on a handheld device. Our prototype presents the visitor with one of a collection of photographs. Each of these photographs was taken facing one wall of a room in a historic house. The visitor changes the viewing perspective (i.e., displays a different photograph) by pressing a button on the device. When the visitor taps on an object in a photograph, the guidebook gives a description of that object, if one is available. Many, but not all, of the objects visible in a given photograph have descriptions. (These objects with associated descriptions are known as *targets*.) Because the historic house environment is complex, many different kinds of objects may be targets. Figure 1 shows a photograph with a number of targets, including a wood panel and a doorway. To help visitors identify targets, the guidebook displays outlines around each target, triggered when the user taps on the photograph but does not "hit" a target.

The visual selection design is motivated by the principles described in [2]. We learned through a combination of observation, informal interviews and professional study [4] that system designs that seem plausible in a museum are not workable in a historic house. Most notably, location-aware systems that use sensors to select content automatically are not feasible in historic houses for a number of reasons (e.g., barriers often prevent visitors from approaching objects). Usability testing of the prototype by thirteen users, conducted prior to the study reported in this paper, confirmed that visual selection is a viable alternative that allows visitors to quickly and easily select objects that interest them.

Figure 1: Electronic guidebook prototype with outlines visible.

The prototype gives the visitor several choices with regard to the presentation of the descriptions. Visitors have the option of seeing a text description of an object or hearing an audio clip with identical content read by a female voice. Visitors can change the choice of textual or audio presentation at any time. Audio clips can be played at a low volume through speakers on the device or through headphones.

The descriptions themselves are typically two or three sentences (40 words) long but do vary in length, with

the audio duration ranging from 3 to 23 seconds. For example, the 20-second description of a portrait of the Duchess of Richmond and Lennox reads:

> "This 17th century portrait shows the Duchess of Richmond and Lennox, about whom Pepys said in his diary, 'Never had a woman more beauty nor less wit.' The portrait was done by Sir Peter Lely, who was the first of the great English portrait painters. Lely created the distinctive look of British portraiture, including the three-quarter pose and the emphasis on beautiful clothing."

METHOD

We conducted an *informal evaluation*, a widely practiced method for studying artifacts that may affect user practices in unanticipated ways [13]. More formal approaches such as controlled experiments, structured interviews, and questionnaire surveys presuppose that the questions of interest are already known. This is inappropriate in cases (such as the introduction of new technology) in which unanticipated uses and results may arise.

In these cases, it is informative to observe people using the new technology in uncontrolled settings that resemble the setting in which the technology will eventually be deployed. User behavior and response to the technology may then be assessed using a variety of methods, including semi-structured interviews that are not limited by the interviewer's preconceived notions but rather encourage the user to speak freely about their experience [11]. The resulting feedback is used to refine the design, after which the procedure can be iterated as necessary. These methods generally involve a smaller number of participants than quantitative studies because they are data-intensive, studying each participant in depth.

In this section, we describe the study participants, the procedure (including the setting) by which we collected the observational data, and the methods by which we analyzed the data.

Participants

The study participants were members of the Xerox PARC community (not necessarily employees), accompanied by friends or relatives with whom they would normally attend a museum. For example, a grandmother attended with her 7-year-old granddaughter and a husband attended with his wife. The visitors comprised a total of seven couples and ranged in age from 7 to over 60 years of age. Two of the couples included a child as a visitor; all other visitors were adults. Eight of the visitors were female and six were male. Visitors were instructed to bring glasses if they used them for reading (to our knowledge one visitor who wore glasses forgot to bring them). One visitor used hearing aids, with which he was slightly hard of hearing. Many of the visitors were non-technical and/or had not previously used a PDA. Two of the visitors had used a previous version of our prototype. Most of the visitors had not previously visited the study site. Half of the visitors described themselves as frequent museum visitors (visiting museums three or more times a year) and half described themselves as infrequent visitors (visiting museums fewer than three times a year).

Procedure

Participants were observed during a private visit to Filoli, a Georgian Revival house in Woodside, California. Each visit consisted of three phases: a partial tour using a paper guidebook, a

partial tour using the electronic guidebook, and an interview.

The visitors went through the first several rooms of the house with a paper guidebook, accompanied by a docent who was available to answer questions. In some cases the docent was one who worked at the house regularly, and in other cases the docent was an escort from the research team (the research escort was always present, even when the regular docent was accompanying the visitors). During this phase, the visitors' comments and conversation were recorded using wireless microphones.

Figure 2: Observation of visitors using the guidebook.

The visitors used the electronic guidebook in the next two rooms of the house. One of these rooms contained security barriers and the other did not. The visitors received brief instructions from the research escort in the use of the guidebook. They were then asked whether they would each like their own guidebook or if they would prefer to share. They were also offered headphones. They were told that they could change their decisions at any time and that the research escort would answer questions about the use of the device (this was rarely necessary). The visitors then toured the two rooms, referencing the electronic guidebooks as desired (the electronic guidebook contained descriptions of 42 objects in the two rooms). The visitors' comments and conversation were recorded using wireless microphones, the visitors were videotaped by a camera placed in a corner of each room (Figure 2), the visitors were directly observed by the research escort, and the visitors' actions in the electronic guidebook were logged by the device for future reference.

After they were finished in the two rooms, the research escort conducted a semi-structured interview. Both visitors in a pair were interviewed simultaneously, which allowed them to respond to each other's ideas and interact with each other as well as with the interviewer. The interviews were videotaped as well as being recorded by the wireless microphones. Some time after the study, two visitors contacted one of the researchers to offer further thoughts on their experience in person and via email.

We intended the interviews to elicit spontaneous responses. Early questions were general, e.g., "How would you describe your experience with the electronic guidebook? What did you like? What did you not like? What would you change?" The interviewer then introduced more specific questions, e.g., "Did you prefer to hear the descriptions in audio mode or text mode? Why?" or "If you were doing this again, would you want to share one [electronic guidebook], or would you want to each have your own?" In many cases, visitors raised these more specific topics before the interviewer introduced them. As one example, many visitors volunteered comments about their opinion of audio versus text before the interviewer introduced the subject.

For a number of reasons, not all visitors answered all questions. For example, the interview guide was refined during

the course of the study, so not all questions were addressed to all visitors. Also, often only a single member of a couple would answer a question directly; to preserve the discussive nature of the interview, the interviewer did not press each visitor for a direct answer.

No time limits were imposed on the visitors during any portion of the study. Visitors spent approximately 20-30 minutes using the electronic guidebooks and approximately 10-15 minutes in the interview. The entire procedure took approximately 75 minutes.

After the visits were completed, we transcribed the comments visitors made during the semi-structured interviews. We then clustered these comments by theme, a technique useful for finding patterns in the values and attitudes held by study subjects [11].

OBSERVED VISITOR BEHAVIOR

In this section, we describe the behavior of the visitors while using the electronic guidebook devices. In the section following this one, we discuss what the visitors reported during the interviews.

Adult/adult vs. adult/child usage patterns

Each member of the five adult-only couples chose to use their own device. (One of these couples began by sharing, but quickly decided that they would each like to use their own device.) Four of these five couples predominantly used audio played through the speakers of the device. Three of these couples conversed frequently, and one conversed quite infrequently. In the other couple, one visitor played audio through the speakers of his device, and his companion listened to audio through headphones. This couple conversed quite infrequently.

One couple consisting of a child and an adult shared a single device during the entire visit, with the child operating the device (playing audio descriptions through the speakers) and the adult looking on and making suggestions. The other couple with a child and an adult (the child's father) began by using separate devices, but gradually evolved to the same shared model as the other child/adult couple. This evolution was initiated by the adult, who gradually stopped using his own device and began to listen to the audio from the child's device. The father requested that the child stay near while the audio was playing so that he could hear it. Further, on occasion the child began to move out of range while the audio was playing, and the father gently restrained him. Both of these couples conversed frequently.

Audio presentation and social interaction

Visitors would frequently stand right next to each other playing different audio clips through the speakers. Visitors had a high tolerance for these overlapping audio clips. Also, visitors often played audio clips multiple times, either to listen to the content again individually, or to share it with their companion. One visitor initially repeated information from text descriptions to her companion, but then realized that playing audio clips was a more effective way to get her companion's attention. She switched to audio mode, saying, "[I] have to play the audio to get your attention, huh?" One couple briefly used two devices to simulate a combined text/audio mode, one member instructing the other, "You get the text and I'll get the audio."

The content of the audio clips often served as a springboard for conversation, either about the objects themselves, or about related topics, e.g., a discussion of how the family spent

Christmas at Filoli led one visitor to remark on an upcoming Thanksgiving celebration of her own.

Effectiveness of a lightweight visual user interface

As suggested by our pre-study usability tests, visual selection proved an effective mechanism, allowing visitors to quickly select objects they were interested in. Visitors selected an average of 37 descriptions, some selecting up to 69 (the latter indicating that visitors viewed or played some of the 42 descriptions more than once). A more detailed critique of the interface is reported elsewhere [3].

While visitors did attend to the device, we did not feel it generally dominated visitor attention. Visitors spent a great deal of time looking at objects in the room, aided by the fact that the guidebook required little attention to operate and by the use of audio as opposed to text descriptions (we discuss this further below). For example, one description mentions detailed carving on a fireplace. When visitors heard this description, they often walked across the large room to inspect the carving, indicating that they were able to effectively transfer their attention from the guidebook to the room.

In summary, the typical adult-adult visit consisted of two adults operating separate devices, playing audio through the speakers on these devices, chatting with each other, and frequently looking at objects in the room. The typical adult-child visit consisted of the child operating a single device, playing audio through the speakers on the device, chatting with the adult, and the adult and child frequently looking at objects in the room.

SELF-REPORTED EXPERIENCES AND VALUES

We believe that the electronic guidebooks were effective because in most cases they helped visitors to strike the desired balance of interaction with all three entities that demanded their attention: the guidebook, the room and its contents, and their companion. Structuring our qualitative analysis in terms of these entities resulted in some interesting insights. In the following subsections, we discuss what visitors wanted from each of these entities and how the guidebook and the delivery mode helped and hindered them. We then discuss how some visitor priorities compete with each other and how visitors deal with these conflicts.

Overall impression

Visitors had a very positive response to their experience with the electronic guidebooks, saying they had “fun” and “liked” them. One visitor later said in email:

> B: “It was one of the richest experiences I have had while visiting a museum.”

Two visitors mentioned establishing a relationship with the guidebooks:

> K: “In fact, you almost establish kind of a relationship or a rapport to the [electronic] guidebook... you feel like it’s kind of, uh, going along with you.”

Visitors also mentioned that they felt they had learned more or spent more time than they would have with other presentation methods:

> J: “I think I probably – in terms of totally self-guided – I probably spent more time with things in the room than I would have if I just were walking through.”

Many visitors directly stated or implied that they would generally choose the electronic guidebook prototype over other presentation methods.

Visitor-guidebook interaction

Two key issues emerged relating to visitor-guidebook interaction. Both relate to the visitors' desire to manage their attention effectively. The first is the degree of control that the guidebook afforded to them with respect to their experience. The second is the amount of attentional demand made on them by the guidebook. In the remainder of this subsection, we discuss these points in more detail, providing illustrative quotes from the visitors.

Control. Visitors highly valued being able to control the information received from the guidebook. They expressed positive feelings towards the electronic guidebook prototype because it gave them a high degree of control.[2] They expressed negative feelings towards other presentation methods that do not provide this degree of control. Visitors also stated that they preferred having their own device (rather than sharing devices) because it helped them maintain control. As we discuss further below, control allows visitors to construct an experience with the desired balance of interaction with multiple entities, e.g., the right amount and types of information from the guidebook, observation of objects in the room, and conversation with a companion.

Visitors wanted to be able to control what the guidebook told them, when it provided the information, and the amount of information they received.

> W: "There may be certain things that don't interest me in which case I don't want to waste my time with them, but other things might intrigue me a lot and I'd like the opportunity to kind of be able to drill down more..."

Many visitors said that they liked the electronic guidebook because it gave them control over which objects they selected, and when (it did not provide control over the level of detail, and several visitors requested this capability). Some visitors spoke of disliking or avoiding linear audio tours or other tours that did not afford this degree of control.

> E: "I usually don't use an audio tour. I don't like to – The thing that's nice about this compared to [a] regular audio tour[3] is that I'm on my own speed, and I hate the audio tours where I'm locked into somebody else's idea about what I want to see."

> S: "The information was the right length. It wasn't too long or too short, I found, you know. Like sometimes, when you take guides – I was in Russia this summer and they just went on and on, you know, for two hours nonstop. You get guided through the – the museums and it's really tough. After a while you tune out... overload."

The value of control further manifested itself in a desire to have exclusive control of the interface, all members of adult-only pairs expressing that they preferred to have their own device (we discuss later how this desire interacts

[2] Naturally, other presentation methods (including the interfaces of many commercial audio guides) do provide this type of control. We discuss these attributes here not because they are unique to our interface but because we are interested in how they interact with other unique attributes of our interface.

[3] We believe that a number of visitors were unfamiliar with random-access audio tours such as the Acoustiguide Inform™, so references such as this one are presumably to linear audio tours.

with their desire to share their experience).

> B: "I would never want to share it with another person. That would drive me insane."
>
> B: "If you had to share it because it was cheaper at the museum and they'd give you one for the same price, then, you could, like, put up with it. But, you know, it's kind of like having a remote control at home on your TV. Do you really like having somebody else poking the channels for you?"
>
> K: "Talk about a personal information system – this is very much a quintessential personal information system."
>
> E: "[K] and I both experience tours in a totally different way and it would be very frustrating to have to share with him."

Attention overhead. Visitors also cared about the degree to which the interface with the guidebook made demands on their attention. When visitors sensed that the guidebook was demanding too much of their attention, they responded negatively: three visitors mentioned that interacting with the guidebook drew their attention away from the room, as we discuss further below.

Conversely, when visitors sensed that the electronic guidebook demanded less of their attention than other presentation methods, they reacted positively. Specific examples tended to share particular themes. Visitors preferred audio because it relaxed demands on their visual attention (we elaborate on this point in the next subsection). Further, several visitors, particularly those who were apparently experienced with keypad-based audio guides, liked the visual selection mechanism. These visitors said that the visual interface was easier to use than numbered labels, helped them confirm that the object they were looking at in the room was the one being described (without a picture, visitors are not always sure they are looking at the right object), and helped them orient themselves in the room. All of these ease-of-use capabilities reduce attentional demands, a design goal of the prototype.

> A: "I liked – It kind of was easier vis- you know, visually clicking on it rather than looking to see that it's number 324 (points to wall) and then [companion laughs and says 'yeah'] 3, 2, 4 (simulates entering number in a keypad)."
>
> A: "[I am about to take a group of children to Underwater World, an aquatic theme park.] This [electronic guidebook] would be, even be easier than... this is going to be a group of 7 and 8-year-olds. It takes a while to tell them, 'ok, you want to learn about the, uh, you know, the leopard shark, you have to do in 1, 2, 4,' you know. Where this, I bet that any 8-year old... would figure it out in, you know, 10 seconds."
>
> K: "We did the audio tour at Alcatraz, for example, and that was very useful. Again, there were some... navigation problems, because we weren't in the right place at the right time. The crowds are large enough so you can't actually move freely... If you had this [electronic guidebook]... you'd have a very strong correlation from the picture as to whether you were actually looking at the right thing."

Visitor-room interaction

Visitors talked about several ways the guidebook impacted their interaction

with the room. First, many visitors said they usually chose audio rather than text descriptions because they could pay more (visual) attention to the room. Second, a few visitors said the interface drew their attention away from the room. Third, a number of different stimuli (some from the electronic guidebook, some not) motivated visitors to examine specific objects in the room. In this subsection, we discuss these issues in turn.

Audio vs. text. The most common comment made about interaction with the room was related to the choice of audio. Many visitors commented on the advantage of being able to look at objects while listening to an audio description. They valued this ability because the audio did not make visual demands on their attention (as compared to text) and because the audio did not force them to context-switch between looking at text and the location.

> W: "When you're reading, you're not looking at the room."
>
> L: "Listening while you're looking at [the object] – that's really value. It's hard to read the text and then go up and look and then read some more text and go up and look..."
>
> B: "What I found – it was very comfortable to not have to read while you were looking. I liked that a lot, you know… it's distracting to have to read the text and then look at the items because you just lose the mood of, you know, experiencing the room…"
>
> D: "So I think the printed guide is more distracting 'cause you're looking at something and then you got to – down – look at the book and – So that's the advantage a lot here."

Some visitors did observe that text had distinct advantages over audio. A few visitors used text rather than audio in specific instances, e.g., to see how a specific word was spelled or to get more detail.

> J: "I kept switching back and forth. So for some of the times I liked just having – Like when I was looking at the birds I liked having the audio and I was hearing the recording – But then there were times when I was trying to underst- when I was trying to pick up more detail and going back to the text – Most of the time I was on the audio so I that I could be looking at something and then go back into the text and read it when I didn't quite process or when I – I don't know – there were times when I wanted to see the text of it."

Distraction from the room. Observations of the visitors and subsequent video analysis suggest that the visitors spent a great deal of time looking at objects in the room (even those who felt the electronic guidebook distracted them to some degree). Some visitors stated that the guidebook interface actually made fewer attentional demands than a paper guidebook that required them to look back and forth or a device that required entry of numbers on a keypad. Nonetheless, a small number of visitors were concerned that the electronic guidebook interface demanded attention and distracted them from the room.

> K: "This really does change the experience, because you end up, it's like you've got a video game in your hand. So... you

> concentrate as much on the device as you do on the surroundings…"
>
> W: "Maybe because it's new, but, you know, this seemed a bit more intrusive... I kind of spent more time futzing with it than looking in the room."

Interestingly, these observations were made by technically-advanced visitors who used PDAs regularly.

Motivation to look at objects. A number of different stimuli motivated visitors to get information about objects in the room. Visitors were strongly motivated by the availability of objects in the guidebook. When visitors tried to choose based on what they saw in the room, they were often disappointed because not all objects have descriptions associated with them.

> E: "[I chose objects from the electronic guidebook because] I very quickly learned that there were only some objects that I could [select]... Whereas if they were all available, like we've talked about, we would like – I think I would just go with, you know, what struck my eye."

Visitors were also motivated by observing objects around them. The following sentiment is supported by many comments that the visitors made while they were using the guidebooks, e.g., a visitor might say, "Where's the ship?" and then look for it in the guidebook.

> V: "You see something that really interests you and you try to find it on here [the electronic guidebook]."

Finally, visitors sometimes overheard information from their companions' device, or their companions shared information with them, as described in the next subsection.

Visitor-visitor interaction

Many visitors (both frequent and infrequent attendees) stated that they had a strong desire to interact socially during the visit. The delivery mode played a large role in social use of the device, creating shared context, letting the visitors know when the other person was available, and providing audio clips that could be shared or overheard. In the remainder of this subsection, we discuss these issues in turn.

Importance of social interaction. In addition to stating that their visit had a social purpose, visitors observed that they behaved differently depending on their companion, indicating a willingness to set aside individual preferences to create a shared environment with a companion.

> S: "It [the type of information preferred] depends on my mood, too. I like both [personal information and historical information], actually. Because I'm also interested in the arts, so it depends on my mood. I think when I'm with my friend – when I'm with you [A], I probably would prefer the his- you know, more the what the house is about, the family history. But if I would come here by myself, I think I would want to learn a little bit more about the – the artwork."
>
> O: "[Previously said he would prefer to use his own device] But there are instances when being able to – you know, if I was going along with my mom or my grandmother, my mom would not have the patience to do this, but me playing and her listening it – to it would – would suit her just fine. So there are instances when I can see sharing it."

The electronic prototype was very successful at facilitating social interaction, and visitors were consciously aware of this:

> G: "A great benefit is the sharing... you share more [using the electronic guidebook] than you do even when you walk together."
>
> J: "I could imagine, like, doing this with my husband or something. There's a social sharing that's happening there that's nice if you were visiting the place with somebody – as opposed to – as opposed to the isolation."

The overall social experience seemed very well received, particularly when compared with traditional presentation methods:

> G: "The sharing is something really nice, because I mean I would not like her to have an earphone and me to have an earphone and we walk around with our different [devices]... then no communication."

Creation of shared context. Shared audio through the speakers facilitated interaction more than paper or headphones, because it created a shared context for the visitors.

> A begins: "I also think the listening versus [reading], although I know, you know, you could obviously do it with a headphone, which we didn't, umm, was kind of even more social to listen to it together. I –"
>
> S interrupts: "Than reading, yes. Exactly. You can make more comments and things, you know. While you are hearing something you can make a comment. While the other person –"
>
> A interjects: "You don't know where they're reading."
>
> S continues: "– is not reading at the same speed, you don't know if they've already caught on to the same speed, so this [shared audio through the speakers] is kind of nice. So you can listen to it together and make a comment, you know, about the – the wine being available or something [referring to a description in the electronic guidebook] (laughs). "
>
> A interjects: "Right, right."
>
> S continues: "No flowers there [referring to a description in the electronic guidebook] (laughs). We wouldn't have – If we read something maybe we wouldn't be at the same place at the same time making a comment, so [this interface is] kind of more social there."

Knowing what others had just learned (either by shared listening or overhearing) provided a basis for direct interaction.

Disclosure of conversational availability. The shared audio not only let visitors know what their companion had heard, it also meant that visitors knew when their companion was available for conversation. By contrast, when a companion is either reading or listening through headphones, visitors do not know when a comment would be welcome or appropriate.

Direct use of guidebook content as conversational material. Audio played through the speakers also allowed the visitors to play audio clips for each other (which they frequently did).

Some visitors also "eavesdropped" on each other's audio.

> J: "The other thing I noticed is, you know, hearing something that was on hers, and going, 'Wait, what was that? That sounded interesting. What did she click on?'"
>
> J: "[I declined headphones because] I wanted to hear what hers [L's device] was saying."

A key aspect of the audio clips used in the prototype was their short duration. They could easily be integrated in conversation, which we believe encouraged visitors to use them routinely; if the segments had been longer, we think they would have been cumbersome and prevented sharing, and therefore, used less frequently.

Competing priorities

As shown above, most visitors took advantage of mechanisms in the guidebook to incorporate the guidebook, their companion, and the room in their experience. To achieve balance among these competing entities, they made a number of complex tradeoffs. For example, people want a shared experience, so they create shared audio environments that allow them to participate with their companions. However, sharing is not so high a priority that they will actually share a device, i.e., that they will give up control and therefore their ability to manage the entities competing for their attention.

These competing desires manifest themselves at a higher level when a visitor chooses whether or not use a presentation method at all. For example, when offered a fully scripted audio tour as opposed to a guidebook, visitors must choose between their desire for control and their desire to get audio rather than textual information. At least in some situations, visitors say the desire for control dominates.

DESIGN IMPLICATIONS

An electronic guidebook must be designed to fit appropriately into a visitor's desired experience. More specifically, designers should be aware that visitors dynamically balance input from multiple attentional entities, and that the guidebook can either support or hinder these efforts.

Our primary findings are that the following principles help visitors balance their attention and therefore maximize their enjoyment of their visit:

- *Provide a way for visitors to share descriptions of objects.* In our study, we specifically found shared "audio spaces" to be highly effective for this purpose.
- *Provide short descriptions of objects.* These descriptions allow the visitors to shift their attention frequently, if needed, and are easily integrated into conversation.
- *Support audio presentation of descriptions.* Visitors strongly prefer audio because it relaxes demands on their visual attention. However, text is sometimes useful for specific purposes, and so some visitors like to be able to switch back and forth between text and audio.
- *Provide random access to information about objects.*
- *Allow visitors to have a personal (unshared) guidebook.* Even though visitors want a shared experience, they want to retain control over their information source to allow them to better coordinate the stimuli demanding their attention.
- *Provide a method of selecting objects visually, or of confirming object selection visually* (as opposed

to using, for example, numbered labels). Visual selection decreases attentional demand and may therefore be less likely to interfere with the flow of the visit than other selection mechanisms.

Of these, the most novel relate to audio use and the visual interface. In particular, the rich and creative use of guidebook audio content in (and around) conversation was enabled by the combination of the shared audio channel and the short, conversation-like audio clip lengths. The demonstrated effectiveness of the visual user interface based on photographic images, which was generally perceived as more lightweight than alternatives such as numeric keypads, is also a new result.

RELATED WORK

Our goal in this project has been to improve visitor experience as measured against essentially self-perceived motivations. This is along the lines of those who assess the quality of learning-oriented leisure activity [7,10]; our evaluation did not focus on, e.g., the learning environment [8] or the aesthetic experience [9] per se. (Of course, one of our design goals was to minimize the amount of work needed to gather information in order to maximize the opportunity to both learn about and appreciate the objects.)

With that in mind, our work can be compared to specific previous efforts in the engineering domain, the museum studies domain, and their intersection. We discuss each in turn.

Electronic guidebook design

Electronic guidebook products include a wide variety of systems from industry leaders Acoustiguide and Antenna Audio, as well as from other vendors such as Ameritech (smARTour), JVC (Audio Guidance System), Organic (eDocent), Visible Interactive (iGo) and Vulcan Northwest (Museum Exhibit Guide). Many research systems have also been built (see, e.g., [5-6,14-18]).

Our system differs primarily from previous systems in its reliance on a lightweight visual interface based on photographic images. It also differs in its use of independent navigation mechanisms for different stages of the object selection task [2]. Finally, it provides the option of either text or audio presentation of identical content.

Museum studies

The attentional entities model used in this paper has a structural similarity to prior models in museum studies. Perhaps the best known of these are Falk and Dierking's Contextual Model of Learning (CML) [7] and its predecessors. The CML describes three contexts (personal, physical, sociocultural) that continuously affect how learners make meaning from their surroundings. Entities in our model resemble the CML's contexts, but our model is concerned with attention management, not learning. It therefore deals with concrete entities, makes the personal "entity" implicit, and includes an entity corresponding to an information source – one without a will of its own and (at least partly) in the visitor's control, but (potentially) active and capable of making demands on the visitor's attention.

The museum learning literature also stresses the importance of control, choice, and communication. Again, these results are in the context of effective learning environments. We make a (slightly) higher-level point: unless the learning tools allow the visitor to satisfy their goals with respect to these aspects of their experience, the visitor may well reject the tool and its associated learning opportunity

outright. (Consider the visitors who refused to use conventional audio tours.)

Some of the relative advantages of audio and text presentation are well understood in the cultural heritage community. For example, Serrell notes that audio allows simultaneous use of eyes and ears but tends to isolate the listener [20]. Acoustiguide's marketing literature states that research "based on a series of surveys at client sites... proves that visitors who access Acoustiguide interpretations learn more about exhibitions – and enjoy them more" [1], and some of this is borne out in the academic literature (e.g., [19]). However, we believe that our observations about short, conversationally compatible audio clips are novel and are not obviously predictable from studies of, e.g., short text labels [20]. We also believe that our observations on the individual use of identical audio and text are new – we are not aware of any other guidebooks that allow the visitor to switch between audio and text presentation in this way. Finally, our results on visitor usage of individually controlled, sharable audio are new as well.

In a study of exhibit label reading, McManus observed high rates of "text echo," inclusion of label text in conversation [12]. The widespread uses of audio sharing in our study (both deliberate and eavesdropped) demonstrate that technology can be used to help visitors introduce label content into conversation directly. McManus also suggested that visitors process – and are inclined to treat – exhibit labels *as conversation*. Again, our findings indicate that technology can bring visitor experience even more in line with this inclination, particularly if the audio descriptions are short enough to easily integrate in existing conversations with companions.

Electronic guidebook studies

A variety of research systems have been deployed but few have resulted in in-depth studies. For example, HyperAudio [14] was deployed but only results of pre-design studies have been reported [17]. Similarly, Hippie [15] was deployed and received initial feedback, but the results of user evaluations are not available [16]; the same is true of Plantations Pathfinder [18]. We are aware of only two electronic guidebook studies resembling ours. In both cases, the methodology was, like ours, based on a combination of interviews, observation, and device activity log analysis. Unlike our work, these studies used a quantitative approach to measure specific aspects or effects of the systems. A University of Salford team evaluated the design of a tablet computer guidebook prototype at the Museum of Science & Industry in Manchester [6], and a Lancaster University team evaluated the design of another tablet computer guidebook prototype in historic Lancaster [5].

Our study and analysis focused on three key aspects that others did not. First, we examined behavioral issues in audio presentation and comparative issues in audio vs. text (as discussed above). Second, we studied sharing behavior of guidebook users, both in terms of devices and content. Finally, our analysis explicitly modelled the guidebook (information source) as a separate attentional entity, which provided an interesting framework for theme identification.

CONCLUSIONS AND FUTURE WORK

Visitors want to construct an experience that incorporates the three entities they value: the guidebook, the room, and their companion(s). To construct this

experience, they attempt to maintain a dynamic balance of the quantity and type of information coming from all three entities to which they want to pay attention.

Our electronic guidebook prototype had a substantive impact on visitors' ability to interact with each other and with the rooms and their contents. Visitors chose guidebook options that facilitated these interactions (and generally ignored those that did not). In this way they were able to create fulfilling experiences, and so visitor response to the electronic guidebook was extremely enthusiastic. Overall, visitors preferred audio played through speakers because it allowed them to interact with both the room and their companions. Short audio clips were particularly desirable since they gave visitors more control and were easily integrated in conversations.

Our current work includes further analysis of the data collected in the course of this study. For example, we are currently conducting a detailed conversation analysis of the video recordings. We are also studying what drives a visitor's inquiry into a particular object (e.g., independent observation of the object in the room, a companion's interest in the object, or the presence of a description in the guidebook).

ACKNOWLEDGMENTS

We are deeply indebted to Tom Rogers and Anne Taylor of Filoli for their generous assistance with this project. We also thank Tom for his perceptive comments on the design of the prototype, Maribeth Back for assistance in recording the audio clips, and Victoria Bellotti, Beki Grinter, Bob Moore, Mark Newman, and Morgan Price for their helpful insights.

REFERENCES

1. "Why an Audio Tour?" Acoustiguide Corp., New York, Dec. 2000. http://www.acoustiguide.com/why/
2. Aoki, P.M. and A. Woodruff, "Improving Electronic Guidebook Interfaces Using a Task-Oriented Design Approach," *Proc. 3rd ACM Conf. on Designing Interactive Sys.*, New York, Aug. 2000, 319-325.
3. Aoki, P.M., A. Hurst and A. Woodruff, "Tap Tips: Lightweight Discovery of Touchscreen Targets," *ACM SIGCHI Conf. on Human Factors in Comp. Sys. Extended Abstracts*, Seattle, WA, Apr. 2001, 237-238.
4. Beard, G.W., *Attingham: The First Forty Years*, Attingham Trust, London, 1991.
5. Cheverst, K., N. Davies, K. Mitchell, A. Friday and C. Efstratiou, "Developing a Context-Aware Electronic Tourist Guide: Some Issues and Experiences," *Proc. ACM SIGCHI Conf. on Human Factors in Comp. Sys.*, den Haag, the Netherlands, Apr. 2000, 17-24.
6. Evans, J.A. and P. Sterry, "Portable Computers & Interactive Museums: A New Paradigm for Interpreting Museum Collections," in *Proc. 5th ICHIM*, Washington, DC, Sep.1999, 93-101.
7. Falk, J.H. and L. D. Dierking, *Learning From Museums*, Altamira Press, Walnut Creek, CA, 2000.
8. Hein, G.E., "The Constructivist Museum," *J. Educ. in Museums 16* (1995), 21-23.

9. Hein, H.S., *The Museum in Transition*, Smithsonian Institution Press, Washington, DC, 2000.
10. Hood, M.G., "Staying Away: Why People Choose Not to Visit Museums," *Museum News 61*, 4 (Apr. 1983), 50-57.
11. Lofland, J. and L.H. Lofland, *Analyzing Social Settings*, Wadsworth, Belmont, CA, 1995.
12. McManus, P.M., "Oh, Yes They Do: How Museum Visitors Read Labels and Interact with Exhibit Texts," *Curator 32*, 3 (1989), 174-189.
13. Newman, W.M. and M.G. Lamming, *Interactive System Design*, Addison Wesley, Reading, MA, 1995.
14. Not, E., D. Petrelli, O. Stock, C. Strapparava and M. Zancanaro, "Person-Oriented Guided Visits in a Physical Museum," *Proc. 4th ICHIM*, Paris, France, Sep. 1997, 69-79.
15. Oppermann, R. and M. Specht, "A Nomadic Information System for Adaptive Exhibition Guidance," *Proc. 5th ICHIM*, Washington, DC, Sep. 1999, 103-109.
16. Oppermann, R. and M. Specht, "A Context-Sensitive Nomadic Exhibition Guide," in *Handheld and Ubiquitous Computing* (Proc. 2nd Int'l Symp., Bristol, UK, Sep. 2000), P. Thomas and H. W. Gellersen (eds.), Springer Verlag, Berlin, 2000, 127-142.
17. Petrelli, D., A. De Angeli and G. Convertino, "A User-Centered Approach to User Modeling," in *User Modeling* (Proc. 7th Int'l Conf., Banff, Alberta, June 1999), J. Kay (ed.), Springer Verlag, Berlin, 1999, 255-264.
18. Rieger, R. and G. Gay, "Using Mobile Computing to Enhance Field Study," in *Computer Supported Collaborative Learning* (Proc. 2nd Conf., Toronto, Ontario, Dec. 1997), R. Hall, N. Miyake and N. Enyedy (eds.), L. Erlbaum & Assoc., Mahwah, NJ, 1997, 215-223.
19. Screven, C. G., "The Effectiveness of Guidance Devices on Visitor Learning," *Curator 18*, 3 (1975), 219-243.
20. Serrell, B., *Exhibit Labels*, Altamira Press, Walnut Creek, CA, 1996.

ABOUT THE AUTHORS

Allison Woodruff is a Member of the Research Staff in the Computer Science Laboratory of the Xerox Palo Alto Research Center. Woodruff holds a Ph.D. in Computer Science from the University of California, Berkeley, an M.S. in Computer Science and an M.A. in Linguistics from the University of California, Davis, and a B.A. in English from California State University, Chico. Her research interests include information visualization, user interface design, and interactive technology for historic houses.

Paul M. Aoki is a Member of the Research Staff in the Computer Science Laboratory of the Xerox Palo Alto Research Center. He holds a B.S. in Electrical Engineering and an M.S. and Ph.D. in Computer Science, all from the University of California, Berkeley. His research interests include implementation techniques for database management systems as well as mobile computing and interaction design.

Amy Hurst is currently a third-year undergraduate at the Georgia Institute of Technology, where she is majoring in Computer Science and is a member of the Contextual Computing Group. Hurst has spent several summers as an intern at the Xerox Palo Alto Research Center. Her research interests include ubiquitous computing.

Margaret H. Szymanski is a Member of the Research Staff in the Systems and Practices Laboratory of the Xerox Palo Alto Research Center. She earned her Ph.D. in Spanish Linguistics from the University of California, Santa Barbara, specializing in the study of language, interaction and social organization. Her current work uses conversation analytic methods to describe how learning and knowledge sharing interactions are achieved in classrooms, museums, and workplaces.