Digital Exhibit Labels in Museums: Promoting Visitor Engagement with Cultural Artifacts

Jessica Roberts[1,2], Amartya Banerjee[1], Annette Hong[1], Steven McGee[3], Michael Horn[1], Matt Matcuk[4]

Northwestern University[1] Evanston, IL, USA

Carnegie Mellon[2] Pittsburgh, PA, USA

The Learning Partnership[3] Western Springs, IL, USA

The Field Museum[4] Chicago, IL, USA

jarobert@andrew.cmu.edu, AmartyaBanerjee2012@u.northwestern.edu, annettehong2018@u.northwestern.edu, mcgee@lponline.net, michael-horn@northwestern.edu, mmatcuk@fieldmusuem.org

ABSTRACT

How can we use interactive displays in museums to help visitors appreciate authentic objects and artifacts that they can't otherwise touch or manipulate? This paper shares results from a design-based research study on the use of interactive displays to help visitors learn about artifacts in an exhibit on the history and culture of China. To explore the potential afforded by these displays, we unobtrusively video recorded 834 museum visitor groups who stopped in front of one collection of objects. Drawing on cognitive models of curiosity, we tested three redesigns of this display, each focusing on a different strategy to spark visitor curiosity, interest, and engagement. To understand the relative effectiveness of these designs, we analyzed visitor interaction and conversation. Our results uncovered significant differences across the conditions suggesting implications for the use of such technology in museums.

Author Keywords

Museums; learning; interactive displays; curiosity.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

A persistent question facing modern natural history museums is how to understand the role of interactive digital technology in the visitor experience. Can technology be used to foster curiosity and engagement around museum collections? Or does it lead to a digital disconnect in which visitors focus more on screens than the objects in front of them? Can technology help enrich conversation and social interaction? Or does it lead to situations in which people are isolated from one another in galleries? Coming to grips with these questions will be critical to the continued relevance of collections-based informal science institutions.

CHI 2018, April 21–26, 2018, Montreal, QC, Canada

https://doi.org/10.1145/3173574.3174197

In this research, we investigate the role of large interactive displays mounted in front of authentic artifacts in a cultural history exhibit (Figure 1). A growing body of research suggests that such digital technology can create engaging learning opportunities in museums [2, 6, 9, 15, 16, 19, 21, 27, 29, 30, 35, 46]. At one end of the spectrum, displays can organize layers of information. Visitors use the display to navigate through content about an exhibit such as text, images, short video clips, and occasional interactive elements. At the other end of the spectrum, displays can create new forms of audience participation with social media, interactive games, augmented reality, and deep zoom interfaces [5, 27, 35, 29, 30, 40, 41]. Despite the broad range of possibilities, we know very little about how to engage audiences around authentic objects that can't be directly touched by visitors and are often mounted behind glass. Furthermore, while interactive displays have been studied as exhibits in their own right, little work has been done on the use of displays to help learners interpret and appreciate authentic artifacts on display (see [22, 34] for exceptions).

This paper shares work from a design-based research project involving a team of learning scientists and computer scientists collaborating with curators and exhibit developers at a large natural history museum. In June 2015, the museum opened a 7,500 square-foot exhibit showcasing 350 artifacts from prehistoric times to early 20th century China, divided into five themed galleries (Figure 5).

Figure 1. A display case with authentic artifacts behind glass and an interactive touchscreen display (bottom).

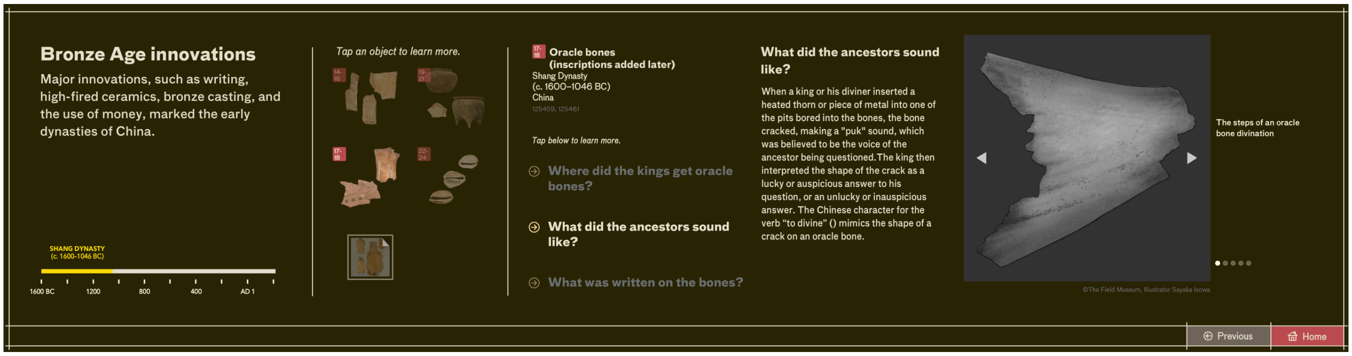

Figure 2. Our first condition (Original) was based on the existing digital rail design, but with slightly modified text to match the other two conditions. Visitors browse information by first selecting from a group of objects and then selecting from one of three questions about the objects.

The exhibit offers a unique opportunity for HCI research as it includes over 45 interactive touchscreen displays called *digital rails* spread throughout the galleries (Figures 1, 2, and 5). Digital rails are designed to serve a similar function to static labels. For example, when visitors approach a display case, they can select from a menu of options to learn more about the various objects in front of them. They can also see information such as maps, timelines, or examples of similar objects not on display. Even though digital rails serve a similar function to that of a traditional text label, they are also interactive computers that could, in theory, provide a range of visitor experiences. The central design tension with these displays is to harness the engagement of interactive media in a way that enhances (rather than detracts from) visitor appreciation of the authentic artifacts on display. In other words, it's important to the museum, its curators, and funders that the artifacts are the stars of the exhibit, not the interactive displays.

To explore the potential afforded by this exhibit, we unobtrusively video recorded 834 museum visitor groups at one of the digital rails midway through the exhibit (circled in red in Figure 5). Drawing on cognitive models of curiosity [26, 28], we tested three redesigns of this rail, each focusing on a different strategy to spark visitor curiosity, interest, and engagement. Our first design (Original) resembled the original digital rail in the exhibit, but with slightly modified content designed to match the other experimental conditions (Figure 2). In this version, visitors navigated a menu by touching pictures of objects to reveal text, images, and interactive content. Our second design (Big Questions) displayed large, provocative questions as a way to pique visitor curiosity and engage them with the display case (Figure 3). As we collected data with these first two designs, we noticed that some of the most interesting visitor conversations took place around slideshow elements that were ordinarily concealed from view—visitors needed to drill down two menu levels to see these images (dealing with the evolution of Chinese characters over time). Therefore, for our third redesign (Timelines), we created interactive timelines based on the existing slideshows of Chinese characters and placed this content front and center. Our intention was to entice visitors to interact with the digital rail first, before presenting text-based content or questions (Figure 4). For our two redesigns (Big Questions and Timelines), we draw on cognitive theories of curiosity that we describe below.

To understand the effectiveness of our redesigns, we analyzed visitor engagement with the exhibit, including whether or not they stopped at the exhibit (capture rate), the amount of time spent at the display case (holding time), the nature of conversations visitors had with one another, and their patterns of interactions with the digital rail. A video camera captured visitors' hands and arms, the digital rail, and audio recordings of conversations. Our results found that the Big Questions and Timeline designs were more effective at attracting visitors to stop in front of the display case and touch the rail than the Original. However, Big Questions was significantly less effective at prompting in-depth visitor conversation, a key indicator of engagement and learning. Furthermore, the pattern of visitor interaction with the rails varied substantially by condition.

Below we review background research and a theoretical framework that informed our designs. We then describe the study design and results including both quantitative outcome measures and qualitative findings that help explain observed differences between the conditions. We conclude with a discussion of design implications that may help guide future work in the creation of interactive digital media for collections-based informal learning institutions such as natural history museums and art museums.

BACKGROUND

Natural History Museums and Authentic Objects

Natural history museums are critical sites for research, education, and public outreach, with numerous studies documenting the attributes that make them valuable sites for informal learning [11, 14, 31, 32]. One of the defining characteristics of natural history museums (as opposed to science centers) is that they are collections-based institutions, and, as such, play a key role in the preservation of our cultural and natural history. Their collections

Figure 3. Our second condition (Big Questions) displays large, provocative questions to pique visitor curiosity and engagement. After tapping on the display, visitors could select from a menu of stories and other objects similar to the original design shown in Figure 2.

typically include "natural" objects such as rocks and fossils as well as cultural objects such as pottery, quilts, and masks. Often, natural history museums are the only source of images, models, or preserved specimens of extinct plant and animal species [39]. These collections also uniquely position natural history museums to engage the public in object-centered learning.

Among scholars who focus on museum learning, there is a general consensus that objects should be framed within a socio-cultural context. Museum visitors talk about objects, listen to others talking about objects, and observe objects with others [1, 7, 10, 11, 14] because they prompt emotive and sensory connections to the political, social, cultural, and historical knowledge that they represent for visitors [43]. Collections-based institutions have a long history of providing resources to support visitor interpretations conversations. This coincides with museum's shifting away from the belief that an object's inherent meaning is contained within the object itself and towards the understanding that an object's meaning lies at the intersection of the object, how it is displayed, and visitors' collective interpretations of the object [10, 11]. While static text labels are the most prevalent support for interpretation that museums provide, we are at a point in time where digital interactives have become a fundamental element of most new exhibitions.

Learning as Conversational Elaboration

To assess the quality of visitor interaction in our study, we build on the notion that museum learning is a social phenomenon rooted in visitors' conversations with one another [25]. Researchers have demonstrated that objects in museum settings support learning through dialogue, with families and friends working together to construct knowledge [3, 7, 14, 25, 42, 47]. This perspective argues that an important way to understand learning in museums is to attend to visitor conversations around exhibits. Important conversational indicators include naming or describing objects, directing attention, asking questions, explaining, and making personal connections [23, 25].

Interactive Surfaces in Museums

With technology improvements, large touch displays have been making their way into museums and other public settings in increasing numbers [e.g. 2, 6, 8, 15, 16, 18, 21, 45]. In the past decade, the field has learned a great deal about how to successfully make use of displays to facilitate collaboration and learning [18, 19, 20, 24, 41, 44]. Early in-the-wild studies have simultaneously highlighted the seductive appeal of large interactive displays for visitors along with serious pitfalls in usability, collaboration, and conflict [21]. Later work has confirmed the difficulty of designing for this context, but has also contributed to a growing collection of design cases that highlighted the potential of this technology in public spaces [2, 6, 18, 35, 45]. Regardless of the research, it is now clear that interactive displays of all shapes and sizes will be commonplace in museums, both as exhibits in their own right and as enhanced replacements of traditional static labels. For this latter use case, the existing literature is very limited. From museum studies, we know about how to design effective static labels [e.g. 13, 38, 43]. However, when it comes to interactive displays designed to help visitors interpret objects, we are in need of empirical evidence, theoretical models, and design principles.

PRELIMINARY STUDY

Before starting the study described below, we conducted observations of visitor behavior in the exhibit hall by observing 49 visitor groups beginning at the entrance of the exhibit and following them as they made their way through the galleries (timing and tracking). We were interested in how long visitor groups spent in the exhibit and what display cases were most popular. As visitors moved through the gallery, we tracked behaviors such as looking at display cases, taking photos, interacting with the digital rails, talking with docents, and interacting with one another. This preliminary data gave us a sense for which display cases were most popular and how long participants generally spent interacting with digital rails. Our results showed that if visitors stopped in front of a display case, they would typically spend a brief amount of time, between 15 and 50 seconds, looking at the objects and sometimes interacting

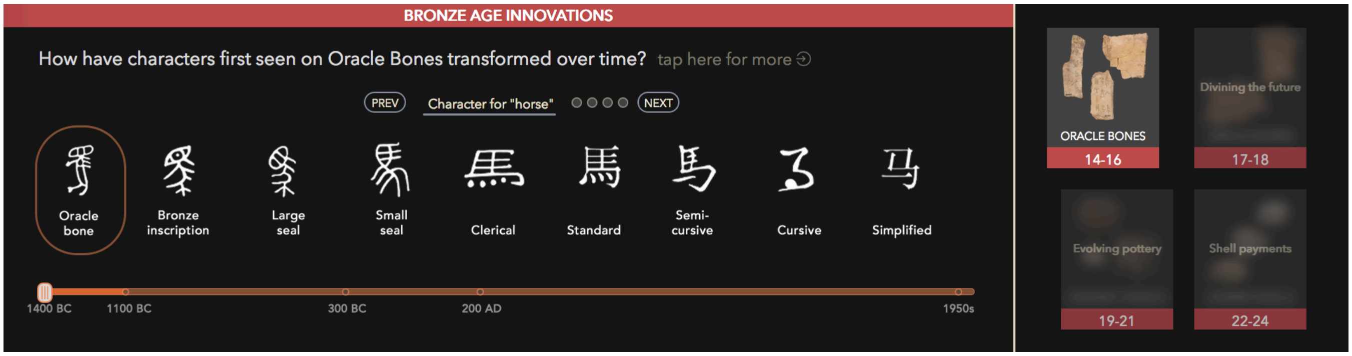

Figure 4. For our third condition (Timelines) we created interactive timelines highlighting the evolution of Chinese characters over time. The oldest known instances of these characters were recorded on oracle bones, examples of which are included in the display case. Visitors can tap on the *more* link to see the same selection of questions and stories as the original design.

with the rail. We found that 30% of the visitors never used any of the digital rails, and the average number of rails used for those who did interact was 4.5 (with a maximum of 25 of the 45 available rails used by a single visitor). We combined these observations with log data from each rail to determine which objects and cases were attracting people and which objects visitors were most curious to read about.

DESIGN

For the next study, we focused on one display case highlighting important Bronze Age innovations from the Shang and Zhou Dynasties, including the emergence of written language, currency systems, and bronze casting. The objects in the case (bits of pottery, cowrie shells, and fragments of oracle bones) are not as visually stunning as other nearby objects, but they have an important story to tell related to the development of social systems in the Bronze Age. However, despite their importance, our preliminary observations showed that this was one of the least-frequently visited display cases in the gallery, and analytics data confirmed this rail was among the lowest in touch counts of the entire exhibit. We therefore focused our design efforts on this case. We sought to spark visitors' curiosity by drawing them in to engage with the content and spurring conversation about the objects.

Our first design (Original, Figure 2) was based on the existing digital rail interface. It consisted of a two-level menu in which visitors would first select a group of objects that interested them by tapping on a picture. This would reveal a selection of three questions about those objects, which we developed to replace topical labels of the existing design. Tapping on a question would then show a short paragraph of text sometimes accompanied by pictures, slideshows, timeline, or an interactive map.

Our second design (Big Questions, Figure 3) was dominated by a picture of a group of objects along with a large, provocative question about those objects. Tapping on the question would reveal a secondary screen with the same text and images as the Original design. The questions shown on the attract screen would rotate every few seconds.

Our third design (Timelines, Figure 4) highlights a slideshow showing the evolution of different Chinese characters over time. The oracle bones displayed in the case are inscribed with the earliest known writing system in China. Through the centuries these original characters have evolved for different purposes and contexts. The original rail design displayed the characters for several words (e.g. "horse," "sun") in different writing styles (from oracle bones to modern cursive and standardized). We built on this content to create interactive timelines that visitors could explore. Visitors could also tap on individual characters to see when that style was (or is) in use. We transformed a second slideshow illustrating the process of oracle bone divination into a similar interactive screen. Visitors could reveal the same story content as the other two designs by pressing the small "tap here for more" link on the any of the interactive screens.

THEORETICAL RATIONALE: CULTIVATING CURIOSITY

In creating our designs, we were influenced by cognitive theories of curiosity. Curiosity is defined as "a desire to know, to see, or to experience that motivates exploratory behavior directed towards the acquisition of new information" [26]. Such behaviors have an obvious appeal for museums with their educational missions to foster visitor engagement and learning, particularly around authentic cultural and historical artifacts. However, while curiosity has been studied with adults in a variety of domains, little work has been done to translate that research into museum contexts to cultivate visitor engagement.

Recent cognitive models have proposed that curiosity has two complementary aspects, *interest* and *information gap*, each rooted in a different neurobiological system [26]. First, curiosity can be thought of as having some degree of negative affect (information gap model) along the lines of tension, frustration, or uncertainty, similar to hunger [28]. As with hunger, feelings of curiosity increase if they are not satisfied but also diminish once information has been obtained. Curiosity increases when one becomes aware of an information gap and has a desire to narrow that gap. Such information could be a fact, the answer to a question,

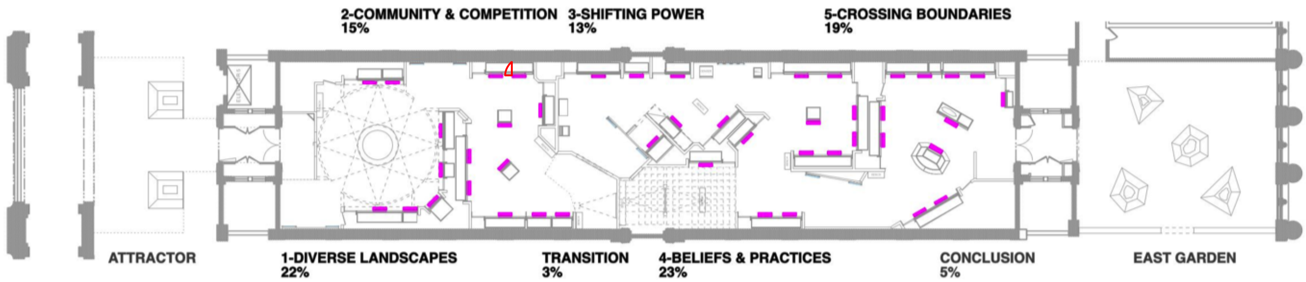

Figure 5. Floorplan of the 7,500 square-foot exhibit space showcasing 350 artifacts divided into five themed galleries. The exhibit features 45 interactive touchscreen displays called digital rails (shown in magenta). Our redesign and study involved the Bronze Age Innovations display case circled in red. This display case is thematically important but infrequently visited.

or the solution to a problem. States of curiosity will intensify as individuals perceive themselves as being close to eliminating their knowledge discrepancy and associated feelings of tension.

Second, curiosity can be thought of having some degree of positive affect when individuals would enjoy discovering something new. In this sense, curiosity is related to the pleasure anticipated from finding out information of a more casual, entertaining, or aesthetically pleasing nature [26]. This positive affect is complementary to the information gap aspect of curiosity.

We believe that our two redesigns (Big Questions and Timelines) appeal to different aspects of curiosity. In the case of Big Questions, posing questions confronts visitors with missing information, which can contribute to information gap curiosity by highlighting what is unknown, but knowable [28]. In contrast, the Timelines design deemphasizes questions and information and instead tries to present an aesthetically pleasing and viscerally engaging experience. In other words, instead of using provocative questions to highlight what the visitors doesn't know, it appeals more to interest and exploratory tendencies. The Timelines design aligns more with frameworks for interactive museum experiences that emphasize engaging visitors on sensory-motor, aesthetic, and emotional levels first, before engaging on a cognitive level [23, 41].

RESEARCH QUESTIONS

Based on these models of curiosity, our research questions concerned whether or not there were differences in visitor engagement, interaction, and conversation among the three designs. We looked at whether or not visitors stopped at the display case, whether or not they interacted with the rail, how long they stayed, and what they talked about. We also analyzed patterns of interaction with the rail across the three conditions. Based on the two interacting theoretical models of curiosity described above, we anticipated that interaction and engagement would be different among the three designs. Further, by explicitly attempting to evoke visitor curiosity, we anticipated that our redesigns (Big Questions and Timelines) would be more successful overall at attracting visitors to the case, at holding their engagement, and at stimulating conversation about content.

RESEARCH METHODS

The three redesigned rails were observed during summer and winter of 2016. Over 40 hours (2,437 minutes) of observations were conducted by three researchers. During observations, a video camera was mounted on a stanchion focusing on the Bronze Age Innovations rail. A sign was posted next to the case indicating that recording was in progress and providing information about the study and contact information for the principle investigators and the Institutional Review Board. No other consent was gathered, and no personally identifiable information was collected in accordance with our IRB protocol. The camera was trained tightly on the video screen to minimize capture of visitors' faces. During observation periods, a member of the research team sat on a bench near the case to record visitor groups entering the "zone", a physical area defined by both physical proximity (i.e. did their path go past the case) and visitor gaze. Visitors walking by the case but not looking at the case were not counted as entering the zone. A timestamp was created when the first visitor in a group entered the zone and when the last member of that group left the zone. Researchers took notes on group interactions.

As noted above, fluid groupings [6] of museum visitors were common around this case, with visitors coming and going at different intervals. Sessions were segmented from the time the first visitor in a group entered the zone to the time the last person in that group left [4]. The entire session was categorized according to the most inclusive category. That is, if a visitor was alone during the entire time in the zone it was coded as a solo session, but if a companion joined for even a few seconds at any point, it was coded as a group session. Groups were, in turn, coded according to whether they talked about the rail or case content (coded "substantive talk") or they didn't talk at all or all conversation was off topic (coded as "no talk"). Visitors speaking a language other than English were coded as "foreign language." The most commonly spoken foreign languages were Spanish and Mandarin. These sessions were translated by native speakers and included in the analysis.

Reading Codes	
Identify	Names or identifies an object (including reading label).
Read Question	Read big question prompt.
Read Text	Read text displayed on the rail.
Interpretation Codes	
Describe	Visually describes a physical object or an image on the display.
Explain	Explains an object or content (summarizing, comparing, providing facts, characterizing).
Remind	Recounts a personal experience related to the object or content (without making a connection).
Commentary	Playful talk and/or opinions about the case content without explaining.
Connection	Makes a connection to another exhibit, object, or culture.
Question Codes	
Question	Asks a question related to the objects or content.
Curiosity Codes	
Information Gap	Statement indicating visitor recognizes something unknown about the objects or information.
Affect	Show affective interest in an object or the rail (including expressions of pleasure or surprise)
Attention	Draw another visitor's attention to an object or something on the rail
Satisfaction	A statement indicating that curiosity has been satisfied ("so that's what it was").

Table 1. Coding scheme used to characterize visitor conversation at the display case.

Sessions in languages other than English, Spanish, or Mandarin were excluded.

Conversation Analysis

It is broadly acknowledged that talk among visitors can greatly impact learning and experience [e.g. 1, 4]. We look at learning talk as a group activity, where ideas can be spoken by one visitor in a single conversational turn or be co-constructed across visitors as they interact with each other and the exhibit. Our coding scheme for analyzing visitor talk was modified from Hohenstein and Tran [17], which was in turn inspired by other work on museum learning [25]. We simplified their original coding scheme to the codes we felt best characterized the conversational turns we were hearing from visitors (see Table 1). We added a set of *curiosity* codes intended to capture instances in which visitors were surprised, confused, interested, or actively curious about any of the objects or content. A single line of the transcript could be coded with at most one code from the Reading, Interpretation, and Question categories and one Curiosity code (see Table 2). Because we were interested in the impact of the rail on visitor engagement, we only coded groups who either touched the rail or referenced the rail directly in some way (e.g. pointing, reading text), resulting in 117 sessions transcribed and coded in this analysis (202 minutes of video).

Idea Unit	Codes
M: See, horse. This is the symbol for horse that they used on the oracle bone.	Explain, *Attention*
M: And this is all...Like (going even). If you wanted to know how to read, you would have to know the different ways of writing	Explain
A: Oh, this is how they read. Oh!	*Satisfaction*
M: This means horse.	Explain
M: See, this is fish.	*Attention*, Identify
M: But this is fish in simplified, and this is fish in standard.	Explain
A: It's so hard!	*Affect*

Table 2. Example of the coding scheme applied to a transcript. An idea unit could be coded with a single *curiosity* code and/or a single code from one of the other three categories.

Two researchers independently segmented the transcripts into *idea units* (lines of conversation), in which dialogue is chunked according to its role in group meaning-making. New idea units indicate a "distinct shift in focus or change in topic or purpose" [36]. The researchers then compared their segmented transcripts and resolved differences. After working together through a training set, the two raters then independently coded 15% of the transcripts to establish inter-coder reliability (agreement 92%).

Participants

In the forty hours of observation we identified 953 visitor sessions. In this analysis, we will only consider sessions in which the attract screen for the design (i.e. the scrolling large text questions in the Big Question design or the interactive timeline in the Timeline design) was displayed on the screen when the visitor approached the rail. This decision ensures that all visitors had similar opportunities to engage with the main features of each design. The screen would reset to its initial state after a certain period of inactivity. Of the 953 sessions, 109 were excluded from the Big Questions design because they did not start on the attract screen, and 5 were excluded from Timeline. An additional 5 sessions were excluded from analysis because they were museum staff, docents, or other visitors aware of the research, leaving 834 sessions analyzed in this dataset.

RESULTS

Our analysis compared the designs based on capture and hold rates, dialogue visitors produced as they used the rail, and patterns of interaction with the touchscreens. Findings in each category are discussed below.

Capture Rate

The two redesigned interfaces used different tactics to attract visitors to the rail content. Big Questions displayed curiosity-inducing questions in a large font that would, at a glance, give visitors a sense for the information they could learn about the artifacts. Timelines brought interactivity to the forefront to encourage visitors to explore. We first sought to determine how these two designs compared to the original interface in attracting visitors (capture rate); that is, what percentage of visitors who entered the zone stopped to look at the case or rail? And, of those, how many stopped to interact with the rail? We defined a stop as a pause where the visitors' feet are not moving. Visitors were marked as interacting with the rail if at any point they used the touchscreen by engaging in a deliberate action such as selecting a story or moving the timeline. Random touching (common with young children) was not counted as an interaction. Some visitors gestured toward the screen or seemed to be looking at it but did not touch; these cases were not counted as interacting.

Condition	In Zone	Stopped	Interacted
Original	291	227 (78%)	93 (32%)
Big Questions	232	186 (80.2%)	99 (42.7%)
Timelines	311	234 (75.2%)	125 (40.2%)

Table 3. Capture rate by condition (solo visitors and groups). Of the visitors who entered the "zone", how many stopped in front of the case and how many interacted with the rail?

Table 3 shows the capture and interaction rates for each design. Based on a chi-square test of independence, the three designs produced no statistically significant differences in capture power (the likelihood that a visitor would stop) ($X^2(2,834) = 1.91$, NS). However, based on a chi-square test of independence, the different designs had a statistically significant effect on the likelihood that a group would interact with the rail ($X^2(2,834) = 7.29$, $p<0.05$). Of the three designs, Original was the least effective at enticing visitors to interact with the rail.

Holding Time

Next, for those groups that stopped in the zone, we looked at the amount of time they spent in the zone for each condition. Session durations were calculated from when the first visitor in a group entered the zone to when the last visitor in a group left, as best as could be determined from the video. Table 4 shows average durations for interactions and non-interactions in each condition. The standard deviation appears in parentheses after the duration. The distribution of holding times can also be seen in Figure 7. Based on a two-way ANOVA, we found, not surprisingly, that visitors who interacted with the touchscreen in any condition stayed significantly longer than those who did not interact ($F(1,833) = 386.9$, $p<0.001$). Otherwise, there were no statistically significant differences across designs ($F(2,832) = 0.03$, NS).

Condition	Interaction	No Interaction
Original (N=227)	79.5 (79.7)	16.0 (15.7)
Big Questions (N=186)	78.4 (61.7)	18.4 (15.2)
Timelines (N=234)	81.7 (67.8)	15.4 (14.5)
Combined	80.0 (69.5)	16.4 (15.2)

Table 4. Average hold times (with standard deviation) in seconds by condition for visitors who interacted with the rail and who did not.

Visitor Talk

We were interested in characterizing the nature of visitor talk in each condition in order to determine how each design supported engagement with the artifacts and rail content. In general discussions at this case were brief—as would be expected in such a large exhibit where visitors on average spent less than 20 minutes traversing all five galleries—but nevertheless some visitors took time to discuss content with their companions. For example, one adult pair in the Original condition discussed the process of divination (how kings prepared, cracked, and "read" their ancestors' words through oracle bones):

M: What'd they sound like?
M: Hm.
M: Wait, did they break the bones?
A: I don't get it.
M: Wait, go to the other steps.
M: Oh, they heated up in order to make, like-
A: And it would make cracks.
M: Wow, that's crazy!
A: Wait, so I don't- but what is that sound?
M: It's probably the, it's the voice of the ancestors
M: So, it's like, certain cracking noises would be like... them talking through the bone, probably?
A: That's how they communicated?
M: I like, this is my ancestor...(inaudible) communicated through.
A: Oh, oh, I get it, okay.

This sort of back and forth information-seeking behavior was a common strategy employed by rail users, particularly in the Original design. The object images and standard identification text (artifact name, approximate age, material, and origin) on the home screen in that design served as an index allowing visitors to easily seek out specific information about an object of interest or browse all available information on that object. The object group "menu" was always available on the left-hand side of the screen if visitors wanted to change course quickly.

In the Big Questions design the menu was still available on the story screens, but the emphasis was the large question displayed on the attract screen. We chose questions we anticipated would spark curiosity by highlighting an information gap: hinting to the visitor that there was some unknown but knowable and interesting piece of information about the objects featured in the case. These questions, we hoped, would spark additional questions, as in this example of a mother and her two kids:

M: "How can a pot shape a city?"
B: Were those, I don't know, I (thought those were teeth).
B: Just for some odd reason I thought that was a tooth.
M: Did you want to learn?
M: It says tap, tap, zoom in, tap here to find out. See there arrow. Pointing there?
M: Ooh, lots of words.
G: (inaudible)
M: (inaudible) half of that.
G: Did they build a big city?
M: They did. They did, they built a...
G: I heard about that.
M: You heard about that?
G: I heard about it. I don't know how I heard about it but I heard about it.

The initial line of this excerpt by the mother is her reading aloud the "big question" on the screen to her children. This initial question prompted them to look for the answer and ask other related questions, e.g. about the size of the city. Here the girl is connecting content from the rail to her past knowledge about sizes of cities in this period of China.

Those connections to prior knowledge are the kinds of conversations museums strive for. The third design, Timelines, was designed to support these kinds of connections because the focal point of the design is the interactive timeline of images, not the transmission of knowledge through text. Visitors therefore could use the interactivity as a jumping off point to find pieces that were most interesting to them, as in the following excerpt of two adult visitors exploring the Timelines condition:

M: ((reading recording sign aloud))
F: Huh!
M: They're doing research on the use of this.
F: Characters. Fish.
F: Small seal, clerical, semi-cursive, cursive, simplified.
F: They're so different!
M: That makes it a little hard.
F: I mean, these look the same. Kind of.
M: This one does too, because that's the (line underneath), see?
F: Here's what it looked like before.
F: Actually, you know it's a funny thing. I went to a talk on the, uh, the evolution-
M: It's the three lines there.
F: - the evolution of, uh, Chinese characters.
M: Mm hm?
F: It is so cool.
M: It's very interesting.
F: Some, you see how old they go.
M: See right here, the line, (inaudible).
M: That looks a lot, seal looks a lot like the simplified right here.
M: Cursive looks cool though.
F: Moons.
F: Look how much they changed.
M: Geeze.
F: Geeze.
F: Fourteen thousand BC
M: (China is really old)
F: Crazy.

As demonstrated in this excerpt, visitor conversations involved explanations, comparisons, reading bits of text, remembering past experiences, and making affective observations ("China is really old"). Though conversations like these are very rich, they are more likely to be centered around the rail itself rather than the objects on display.

To systematically compare visitor conversations across conditions, we used the scheme in Table 1 to code the transcripts of all sessions involving groups of visitors (2 or more people) who interacted with the rail and spoke English, Mandarin, or Spanish. We first looked at the overall likelihood that visitors had any substantive talk in their session. Table 5 shows the number of visitor groups who engaged in substantive talk if they stopped at the rail. Of those, we also counted the number of visitor groups who engaged in substantive talk if they interacted with the rail. In neither case was there a statistically significant difference based on chi-square tests of independence for whether groups stopped ($X^2(2,316) = 1.3$, NS) or interacted with the rails ($X^2(2,182) = 1.5$, NS).

Condition	Stopped and Talked	Interacted and Talked
Original	60 / 110 (54.5%)	33 / 53 (62.3%)
Big Questions	55 / 99 (55.5%)	34 / 57 (59.6%)
Timelines	66 / 107 (61.7%)	50 / 72 (69.4%)

Table 5. Proportion of visitor groups (2 or more people) who engaged in substantive talk if they stopped in front of the rail and if they interacted with the rail.

Though there wasn't a difference among conditions in likelihood of talk, we wondered if any design seemed to support more talk overall or of a particular kind. We therefore looked at the frequency of talk code use across the conditions. To conduct this analysis, we collapsed the codes into four categories of talk: *Interpretive Talk* category consisted of the Describe, Explain, Remind, Commentary, and Connection codes and represented visitor attempts to make sense of the objects or content; the *Reading* category consisted of the Identify, Read Question, and Read Text codes, in which visitors were not generating content but were sharing text from the rail aloud with each other; the *Curiosity* category consisted of all four curiosity codes, indicating when visitors wanted to know something or found something they found interesting; and the Questions category consisted of the single Question code, which encompassed all types of questions ranging from short single answer information seeking questions (e.g. "What is

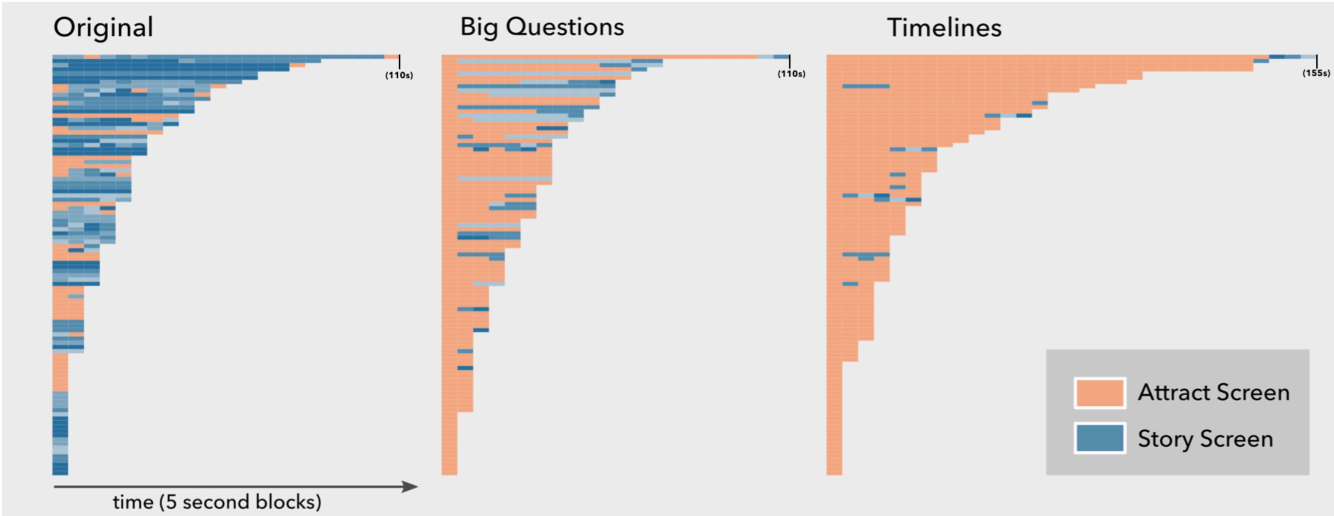

Figure 7. A random subsample of 100 visitor sessions in each condition sorted by length. Each horizontal row shows one visitor session, and each vertical column is one 5-second window. Shades of blue represent the different story screens, while orange represents attract screens.

this?") to more open-ended speculative questions, (e.g. "Why wouldn't the king just say whatever he wanted?").

We then looked at the average number of times each category was used per session in each condition (see Figure 6). Since the talk codes are counts for each session, a Poisson regression was run to estimate differences in the frequency of each talk code based on design. There was a statistically significant difference in the frequency of Interpretive and Curiosity codes. Groups in the Big Questions condition less frequently engaged in Interpretive talk ($ß = -0.56$, $p < 0.001$) and Curiosity talk ($ß = -0.89$, $p < 0.001$) than in the other two designs. There was no statistically significant difference in the frequency of Questions asked or Reading of text on the screen.

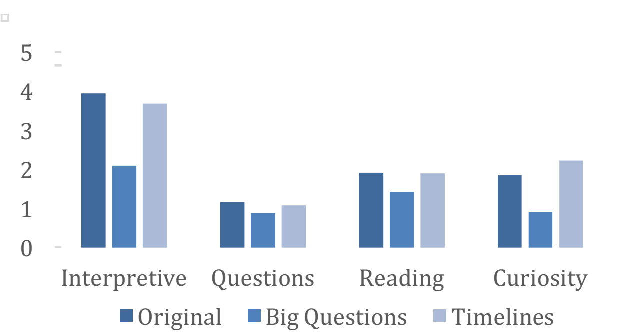

Figure 6. Average number of idea units per session that were coded in four different categories. Grouped by condition.

Rail Interaction

In addition to analyzing visitor talk, we also looked at patterns of visitor interaction with the digital rail. We began by dividing the video in to 5-second windows for each visitor session where someone stopped in front of the rail (227 sessions in Original, 186 in Big Questions, and 234 in Timelines). We then recorded which screen visitors were viewing at the beginning of that 5-second window. We classified screen content into two categories, attract screen and story screen. For the Original design, the attract screen was visually very similar to the story screens with half the screen content persisting across visitor interactions. The Big Question and Timelines designs had more distinct attract and story screens. Figure 7 shows a random subsample of 100 visitor sessions in each condition sorted by length. Each horizontal row shows one visitor session, and each vertical column is one 5-second window. Shades of blue represent the different story screens, while orange represents attract screens. Note that in the Timelines condition, visitors could interact with the timeline without transitioning between screens, so this would show up as longer orange segments.

The left side of Figure 7 shows that visitors using the Original design tended to transition relatively quickly from story screen to story screen, giving them opportunities to see more about the objects on display. In contrast, in the Big Questions condition (Figure 7, middle) visitors first saw the attract screen with the provocative question and were less likely to subsequently transition between many different stories. Their explorations were typically limited to one or two story screens. Finally, in the Timelines condition (Figure 7, right), visitors spent almost all of their time interacting with the timelines and slideshows, which were also the attract screens for this condition, and much less time on the story screens containing information about objects in the display case.

DISCUSSION AND IMPLICATIONS

In this study, we tested three design variations for an interactive digital reading rail as part of a large cultural exhibit at a natural history museum. The first design (Original) used a standard drill-down menu interface now common in museums. Visitors could select pictures of objects to learn more about them. This model might be thought of as an information-on-demand approach. This approach was the least effective at enticing visitors to interact; however, once groups engaged with the rail, they tended to browse through a number of different stories and had relatively in-depth conversations.

For our second design, we built on a cognitive model of curiosity (knowledge gap) characterized by feelings of tension, frustration, and uncertainty around not knowing something—and the satisfaction one gets from sating this feeling. We explicitly designed questions to pique visitor curiosity by highlighting unknown but obtainable information. This approach was somewhat successful in that visitors were more likely to stop in front of the display case (not significant) and interact with the rail (significant) in this condition than the Original. However, in the information gap model, feelings of curiosity diminish once the desired information has been obtained. As there were no additional curiosity inducing features to prompt visitors to pursue further information, this might explain why visitors were only likely to view one or two stories on the rail.

In addition, visitors engaged in significantly less interpretive talk (describing, explaining, commenting, and making connections) and curiosity talk (acknowledging an information gap, expressing affect, directing attention, and noting satisfaction that curiosity has been satisfied). This result was somewhat surprising in that our curiosity codes had been developed with the intention of detecting information gap curiosity. It's possible that once visitors get into a question/answer mode, they are more likely to see the museum as an authority figure that dispenses knowledgeable information and less likely to engage in self-directed meaning making and interpretation. This would align with past research on science museums that distinguishes between "planned discovery" exhibits and more open-ended "active prolonged engagement" exhibits [23]. With planned discovery, visitors spend relatively less time engaging and tend to move on once the pre-designed "discovery" has been made.

Our third design aligned more with an interest model of curiosity in which exploratory activity is initiated from the pleasure of finding out new information that is more casual, entertaining, or aesthetically pleasing in nature. This design was also more effective at enticing visitors to interact than the Original design, and it was significantly better than Big Questions in eliciting Interpretive and Curiosity talk (although this condition did not perform any better than the Original design in terms of visitor talk). It also did not seem to encourage visitors to explore more about the objects—for example, visitors seemed to rarely make the connection between the evolution of Chinese characters and the actual oracle bones in the display case.

This work highlights the design tensions in using digital media to pique curiosity about museum objects and enable conversations about the meaning of objects. On the one hand, while bold, provocative questions on interactive exhibit displays might draw visitors in, this strategy can also backfire, leading to relatively impoverished conversation. On the other hand, offering less text-heavy and more viscerally engaging displays can entice visitors to interact (perhaps appealing to interest-based curiosity), but this approach can also train visitors' attention on the displays themselves rather than on the objects they were meant to interpret.

To resolve this conundrum, museums might design displays that use the big questions approach to appeal to knowledge gap curiosity, but also design in layers of content that induce follow-up questions in visitors' minds, thus prolonging engagement. However, there is still a risk that leading with questions will have a deleterious effect on visitors' self-directed meaning making (as we saw in our data). Another possibility is to design displays that appeal more to interest-based curiosity (using less text and more viscerally engaging media), but then design in much more explicit references to the objects themselves. These references could even take the form of knowledge gap questions that visitors might be more amenable to after engaging in more open, self-directed inquiry beforehand. Our data also suggest that the information-on-demand approach can work well, especially when visitors are already curious enough about the objects themselves to have a desire to learn more.

ACKNOWLEDGMENTS

The authors would like to Jaap Hoogstraten, Deborah Bekken, Taylor Peterson, Patience Baach, Christina Yang, Eureka Foong, and Tyler Lindow for their assistance with interface design, data collection, and coding. This research was supported by grant IIS-1550051 from the National Science Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

REFERENCES

1. Sue Allen. 2002. Looking for learning in visitor talk: A methodological exploration. In Learning Conversations in Museums, G Leinhardt, Kevin Crowley and K Knutson (eds.). Learning conversations in museums, 259–303. http://doi.org/10.1207/S1532690XCI1704_1
2. Antle, A.N., Tanenbaum, J., Seaborn K., Bevans, A., & Wang, S. (2011). Futura: Design for collaborative learning and game play on a multi-touch digital tabletop. In Proc. Tangible, Embedded and Embodied Interaction (TEI'11), ACM Press, 93-100.
3. Ash, D. (2002) "Negotiation of Biological Conversations in Informal Learning Settings," in G. Leinhardt, K. Crowley and K. Knutson (eds) Learning Conversations: Explanation and Identity in Museums, pp. 371–400. Mahwah, NJ: Laurence Erlbaum Associates.

4. Leslie Atkins, Lisanne Velez, David Goudy, & Kevin N. Dunbar. 2009. The unintended effects of interactive objects and labels in the science museum. Science Education, 93(1), 161-184.
5. Beheshti, E., Kim, D., Ecanow, G., & Horn, M. (2017). Looking inside the wires. Understanding museum visitor learning with an augmented circuit exhibit. ACM Conference on Human Factors in Computing Systems (CHI'17).
6. Block, F., Hammerman, J., Horn, M.S., Phillips, B.C., Evans, E.M., Diamond, J., Shen, C. (2015). Fluid grouping: Quantifying group engagement around interactive tabletop exhibits in the wild. ACM Conference on Human Factors in Computing Systems (CHI'15), ACM Press, 867-876.
7. Borun, M. (2002). Object-Based Learning and Family Groups. In S.G. Paris, Perspectives on Object-Centered Learning in Museums (1st ed). Mahwah, NJ: Erlbaum.
8. DiAngelo, S., Pollock, D.H., & Horn, M.S. (2015). Fishing with Friends: Tabletop games to raise environmental awareness in aquariums. Interaction Design and Children (IDC'15), 29-38, ACM Press.
9. Catherine Eberbach and Kevin Crowley. 2005. From living to virtual: Learning from museum objects. Curator: The Museum Journal 48, 3: 317–338. http://doi.org/10.1111/j.2151-6952.2005.tb00175.x
10. Evans, E., Mull, M., & Poling, D. (2002). The authentic object? A child's-eye view. Perspectives On Object-Centered Learning In Museums, 55--77.
11. John H Falk and Lynn D Dierking. 2000. Learning from museums: Visitor experiences and the making of meaning. Altamira Press.
12. Fosh, L., Lorenz, K., Benford, S., & Koleva, B. (2015). Personal and social? Designing personalised experiences for groups in museums. MW Museums and the Web. Chicago.
13. Susanna E Hapgood and Annemarie Sullivan Palincsar. 2002. Fostering an Investigatory Stance: Using Text to Mediate Inquiry with Museum Objects. In Perspectives on Object-Centered Learning in Museums, Scott G Paris (ed.).
14. Hein, G. (1998). Learning in the museum (1st ed.). London: Routledge.
15. Hinrichs, U., & Carpendale, S. (2011). Gestures in the Wild: Studying Multi-Touch Gesture Sequences on Interactive Tabletop Exhibits. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), 3023-3032.
16. Hinrichs, U., Schmidt, H., & Carpendale, S. (2008). EMDialog: Bringing Information Visualization into the Museum. IEEE Transactions on Visualization and Computer Graphics, 14(6), 1181-1188.
17. Hohenstein, Jill, and Lynn Uyen Tran. "Use of questions in exhibit labels to generate explanatory conversation among science museum visitors." International Journal of Science Education 29.12 (2007): 1557-1580.
18. Michael Horn, Zeina Atrash Leong, Florian Block, et al. 2012. Of BATs and APEs: an interactive tabletop game for natural history museums. the 2012 ACM annual conference: 2059–2068. http://doi.org/10.1145/2207676.2208355
19. Horn, M. S., Phillips, B. C., Evans, E. M., Block, F., Diamond, J., & Shen, C. (2016). Visualizing biological data in museums: Visitor learning with an interactive tree of life exhibit. Journal of Research in Science Teaching, 53(6), 895-918.
20. Eva Hornecker, Paul Marshall, and Yvonne Rogers. 2007. From entry to access: how shareability comes about. http://doi.org/10.1145/1314161.1314191
21. Eva Hornecker. 2008. "I don't understand it either, but it is cool" – Visitor Interactions with a Multi-Touch Table in a Museum. Proc. IEEE Tabletop: 1–8.
22. Eva Hornecker. 2010. Interactions around a contextually embedded system. http://doi.org/10.1145/1709886.1709916
23. Humphrey, T. and Gutwill, J. Fostering active prolonged engagement: The art of creating APE exhibits. Exploratorium (2005).
24. Ichino, J., Isoda, K., Ueda, T., & Satoh, R. (2016, February). Effects of the display angle on social behaviors of the people around the display: A field study at a museum. In Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (pp. 26-37). ACM.
25. Gaea Leinhardt and Kevin Crowley. 2002. Objects of learning, objects of talk: Changing minds in museums. In Perspectives on Object-Centered Learning in Museums.
26. Jordan Litman. 2005. Curiosity and the pleasures of learning: Wanting and liking new information. Cognition & Emotion 19, 6: 793–814. http://doi.org/10.1080/02699930541000101
27. Marti Louw & Kevin Crowley. 2013. New ways of looking and learning in natural history museums: The use of gigapixel imaging to bring science and publics together. Curator: The Museum Journal, 56(1), 87-104.
28. George Loewenstein. 1994. The psychology of curiosity: A review and reinterpretation. Psychological

bulletin 116, 1: 75–98. http://doi.org/10.1037/0033-2909.116.1.75

29. Ma, J., Liao, I., Ma, K. L., & Frazier, J. (2012). Living liquid: Design and evaluation of an exploratory visualization tool for museum visitors. IEEE Transactions on Visualization and Computer Graphics, 18(12), 2799-2808.

30. Meisner, R., vom Lehn, D., Heath, C., Burch, A., Gammon, B., & Reisman, M. (2007). Exhibiting Performance: Co-participation in science centres and museums. Int. Journal of Science Education, 29(12), 1531-1555.

31. National Research Council. (2009). Learning Science in Informal Environments: People, Places, and Pursuits. Committee on Learning Science in Informal Environments. Philip Bell, Bruce Lewenstein, Andrew W. Shouse, and Michael A. Feder, editors. Washington, DC: The National Academies Press.

32. Scon G. Paris. 1997. Situated Motivation and Informal Learning. The Journal of Museum Education 22, 2/3: 22–27.

33. Poling, Devereaux A., and E. Margaret Evans. "Why do birds of a feather flock together? Developmental change in the use of multiple explanations: Intention, teleology and essentialism." British Journal of Developmental Psychology 20.1 (2002): 89-112.

34. Quintana, C (2012). “Pervasive science: Using mobile devices and the cloud to support science education anytime, anyplace”. Interactions, 19, 4, 76-80

35. Jessica Roberts, Leilah Lyons, Francesco Cafaro, and Rebecca Eydt. 2014. Interpreting Data from Within: Supporting Human-Data Interaction in Museum Exhibits Through Personalization. In Proceedings of the 13th International conference on Interaction Design & Children. Aarhus, Denmark.

36. Roberts, Jessica, and Leilah Lyons. "Scoring Qualitative Informal Learning Dialogue: The SQuILD Method for Measuring Museum Learning Talk." Philadelphia, PA: International Society of the Learning Sciences., 2017.

37. Rowe, J. P., Lobene, E. V., Mott, B. W., & Lester, J. C. (2014). Play in the museum: Designing game-based learning environments for informal education settings. In FDG.

38. Beverly Serrell. 2015. Exhibit labels: An interpretive approach. Rowman & Littlefield.

39. Shaffer, H., Fisher, R., & Davidson, C. (1998). The role of natural history collections in documenting species declines. Trends in Ecology & Evolution, 13(1), 27--30.

40. Nina Simon (2010). The participatory museum. Museum 2.0.

41. Scott S Snibbe and Hayes S Raffle. 2009. Social immersive media: pursuing best practices for multi-user interactive camera/projector exhibits. Proceedings of the Conference on Human Factors in Computing Systems: 1447–1456. http://doi.org/10.1145/1518701.1518920

42. Stevens, R., & Hall, R. (1997). Seeing tornado: How video traces mediate visitor understandings of (natural?) phenomena in a science museum. Science Education, 81(6), 735--747.

43. Tran, L. (2008). The work of science museum educators. Museum Management and Curatorship, 23(2), 135--153.

44. Xambó, Anna, et al. "Let's jam the reactable: Peer learning during musical improvisation with a tabletop tangible interface." ACM Transactions on Computer-Human Interaction (TOCHI) 20.6 (2013): 36.

45. Xambó, A., Hornecker, E., Marshall, P., Jordà, S., Dobbyn, C., & Laney, R. (2016). Exploring social interaction with a tangible music interface. Interacting with Computers, 29(2), 248-270.

46. Yoon, S. A., Elinich, K., Wang, J., Steinmeier, C., & Tucker, S. (2012). Using augmented reality and knowledge-building scaffolds to improve learning in a science museum. International Journal of Computer-Supported Collaborative Learning, 7(4), 519-541.

47. Zimmerman, H., Reeve, S., & Bell, P. (2008). Distributed expertise in a science center: Social and intellectual role-taking by families. Journal of Museum Education, 33(2), 143--152.