Open Conversation
An IPS-Roundware Application

Formative Evaluation

Joyce Ma and Melissa Zabel

August 2015

THIS IS **NOT** A DEFINITIVE FINAL REPORT

FORMATIVE evaluation studies like this one often:

- **are conducted quickly**, which may mean
  - small sample sizes
  - expedited analyses
  - brief reports

- **look at an earlier version** of the exhibit/program, which may mean
  - a focus on problems and solutions, rather than successes
  - a change in form or title of the final exhibit/program

Formative Evaluation

Open Conversation

An IPS-Roundware Application

Joyce Ma and Melissa Zabel

August 2015

BACKGROUND

Roundware is an open-source framework that collects, stores, and delivers audio content. Created by Halsey Burgund in 2007 for an artist commission with the Aldrich Contemporary Art Museum, it has since developed into a location-sensitive audio platform that can weave comments and stories from museum staff and visitors into a sound tapestry that visitors listen to and contribute to, with a smartphone, as they move through a museum gallery or space. To date, Roundware has been used predominantly outdoors; examples include the deCordova Sculpture Park and Museum and the de Young Fine Arts Museum in San Francisco. In these applications, visitors walk through outdoor sculpture gardens while location-specific stories, commentary and music play on their mobile phones. Prior evaluation results point to the potential of these Roundware apps to encourage visitors to pay closer attention to the art they encounter (Girardeau, Beaman, Pressley, & Reinier, 2015).

As part of the Exploratorium's Indoor Positioning System (IPS) project, we sought to use Roundware to develop a crowd-sourced, location-tagged audio app, called Exploratorium Voices, or Open Conversation, that visitors could use on smartphones to listen to short comments from staff, experts and other visitors and to leave their own comments for others to hear. However, unlike prior Roundware apps, the IPS-Roundware app would use location (longitude and latitude) data derived from a Wi-Fi indoor positioning system instead of the Global Positioning System (GPS) to tag audio recordings and determine where a visitor is to play recordings left at or near that location.

The IPS Project's work with Roundware served two purposes: First, as a proof-of-concept, it identified and addressed technical issues in integrating Roundware with a Wi-Fi based indoor positioning system, and, therefore, represents an important step in the development of Roundware as a flexible platform that can accommodate a number of positioning systems. Second, formative evaluation with visitors gave us an initial glimpse at the nature of visitors' experiences with an indoor mobile app developed with Roundware for a science museum known for its hands-on exhibits. The work documented here focuses on the latter purpose.

EVALUATION PURPOSE

This study looked at visitors' experience with the IPS- Roundware app, Exploratorium Voices: Open Conversation, and examined the following questions:

- What did visitors do with the app?
- What aspects of the experience did visitors find worthwhile and not worthwhile?
- Did the comments they listened to match the exhibit they were at? What did they do when there was a mismatch? How did mismatches affect the visitor experience?

The findings serve to identify the potential and challenges of the IPS-Roundware platform in supporting a location-aware visitor application inside a science museum gallery.

METHOD

The Setting

The evaluation was conducted in the Observatory Gallery of the Exploratorium. (See Figure 1.) The Observatory is the smallest gallery in the Exploratorium, measuring 3500 sq. ft., and containing approximately 26 exhibits, all of which focus on some aspect of the local environment. Ceiling to floor glass windows give visitors a panoramic view of the San Francisco Bay and of the city. Visitors can also access an outdoor terrace from the Observatory, but this terrace was not covered by the IPS and, therefore, not considered in this study.

Figure 1. Inside the Observatory Gallery

a. Looking towards the city *b. Looking towards the bay*

The IPS hardware inside the Observatory consisted of five Wi-Fi Access Points (APs) mounted on beams on the perimeter of the gallery, close to the ceiling. To generate an exhibit map for the Observatory, we followed the ground truth protocol described in Appendix A, which allowed us to derive the position of each exhibit, according to the IPS. The map is shown in Figure 2.

Figure 2. Exhibit location map for the Observatory gallery. These are the location of each exhibit as seen by the indoor positioning system and does not indicate its absolute, physical location.

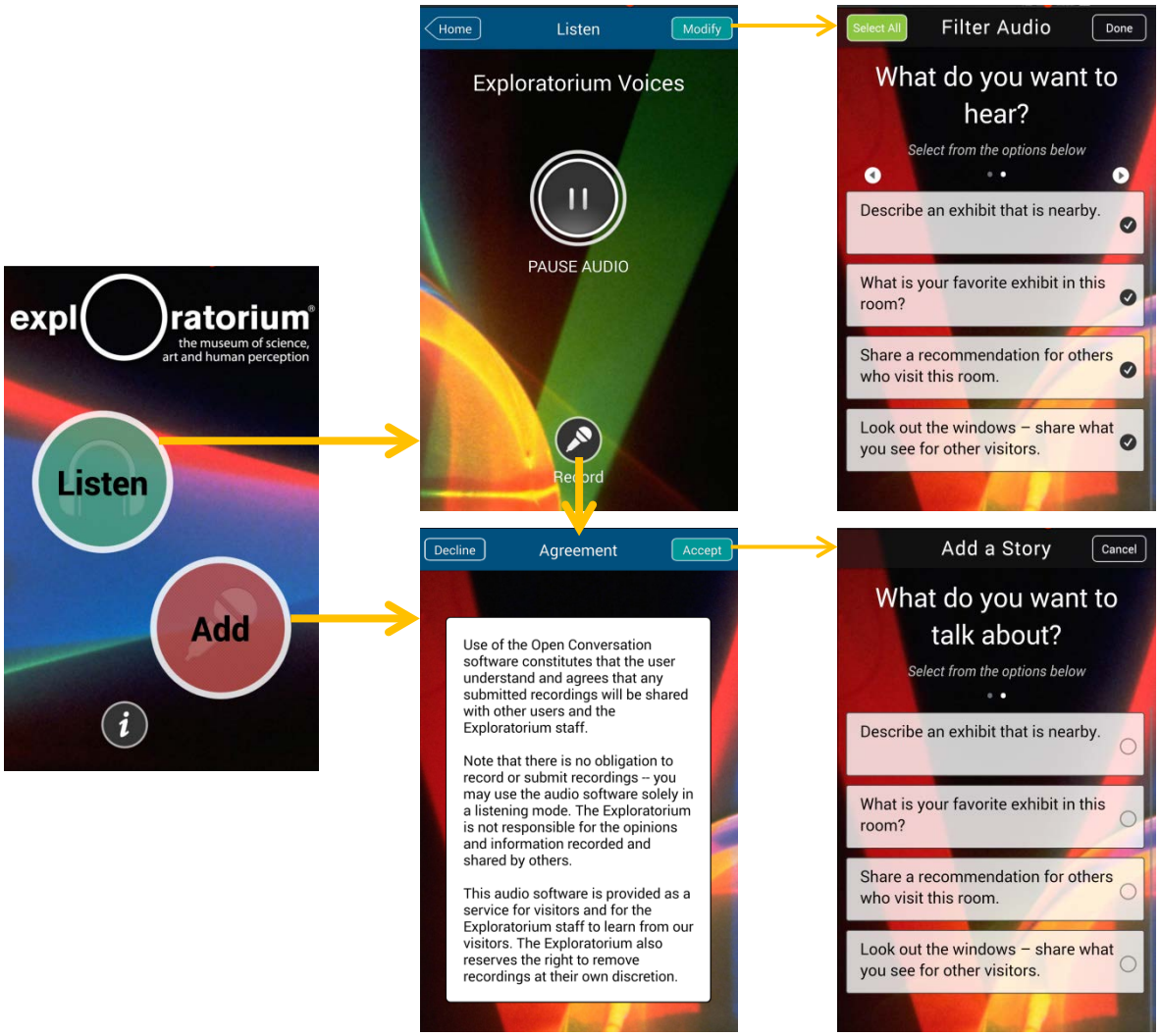
Ground truth tests of the IPS yielded a mean resolution of two meters across exhibits, but the resolution, in fact, ranged from 0.77m to 4.4m depending on exhibit placement, the position of the APs, and the footprint the exhibit. We, therefore, knew that we would not have exhibit-level resolution for this formative evaluation, and that Open Conversation may play audio that were mismatched to the exhibit a visitor was using.

The Application – Open Conversation

The custom-build IPS-Roundware app, Open Conversation, allowed visitors to listen to a continuous audio stream of comments and music during their visit inside the Observatory gallery. Figure 3 shows screenshots of the app that highlight the main tasks it supports. A comment would play when a visitor was at or near a location where that comment was first recorded and tagged by the IPS. Roundware also allowed a visitor to add a comment, which would be location-tagged according to the longitude and

latitude derived by the IPS. Although we had originally planned to develop an app that could be downloaded by visitors to use on their own mobile phones, because of proprietary restrictions on the IPS SDK, Open Conversation was only installed on special smartphones, which we made available to study participants.

Figure 3. Screenshots of the IPS-Roundware app, Open Conversation.



The app was seeded with a few dozen remarks from visitors, staff, artists and scientist, each lasting no more than 30 seconds. Many of these remarks pertained to a specific exhibit, but some were about the gallery itself or about the outside landscape visible through the windows of the Observatory. Table 1 tallies the different types of audio recordings included in Open Conversation. Because many of the exhibit-specific recordings were made before the final fine tuning of the IPS, we had to manually change the location information for each audio file in the Roundware database and placed each exhibit-specific comment at its IPS location.

Table 1. Audio recordings made by visitors, staff, scientists and artists before formative evaluation

Visitors		
Topic	Count	Sample Comment
Exhibit	10	I always wonder how they make 3D models. If you go to the Visualizing the Bay Area, there's one there that's all in white plastic. As a child, we used to make them out of paper mache. But, this one, if you look at the display next to it, explains how they use a CNC router to make the counters look very realistic. So go take a look and feel along it just how realistic it is.
Gallery	0	N/A
Outside Landscape	2	I'm watching a pelican skim the edge of the Bay on this bright sunny day. I'm drawn to the water. I want to be close to the water, and I am enjoying observing it here. I wish I could touch it. I wish I could smell it. I wish I could hear it from here in the Observatory. But instead there is glass between me and the water, and all I can do is to try and understand…
Scientists		
Topic	Count	Sample Comment
Exhibit	4	I'm standing by an exhibit named Tidal Ribbon. It's the favorite exhibit for people who are blind. It's this wonderful, tactile wave in clear plastic that records the tidal heights for the entire year one day at a time. It's great to see the flow of high and low tides from day to day and from month to month and feel them with your hands.
Gallery	1	I'm in the Observatory in the afternoon, and the sun is shining through the west-facing windows. Looking at the ceiling, I can see light reflected onto the ceiling, and I see that the floor is glowing in the sunlight. I want to trace the flow of the sunlight from the sun, through the window, to the floor, and onto the ceiling. I'm standing here and looking at my shadow moving on the ceiling.
Outside Landscape	0	N/A
Artists		
Topic	Count	Sample Comment
Exhibit	10	All right, I'm standing outside, or right next to, Tide Column. The water level shows the current height of the tide, anywhere from 0-6 feet, so I guess that's how much the tide varies here.
Gallery	1	What I want to talk about is the incredible view we have from this gallery. Its wrap-around glass windows, and you can look out on the city as I am right now with the Transamerica Building and Coit Tower and Telegraph Hill. Or if you go to the other side of the gallery, you can look on the water and the bridge and the island. It's beautiful.
Outside Landscape	1	One thought I had is, y'know, these giant clipper ships that used to fill this water, y'know. They come in around the Bay, and come in so there's sort of ghosts in my head of these ships that I've actually never seen. And then there's actual ghost ships that are the story about the fact that the Chinese would send the bodies home. When people died in Chinatown, they would actually send them all the way back to China for burial, so there were these death ships that were essentially carrying caskets of dead bodies back regularly throughout the 19th century. I

don't know how long that went on for precisely.

Staff		
Topic	Count	Sample Comment
Exhibit	16	I am now standing in front of Picturing Place, so this is a favorite exhibit of mine because I had a hand in developing it. It came from a little study that I did looking at what people found interesting around the neighborhood, just taking things from what people had publicly posted on Flickr.
Gallery	1	We're sitting in the open area of the Observatory, internally called the postage stamp because it's on the top area of the building, and it seems like the very end of the Earth almost. You walk in and then there's all these windows that you can look out to the Bay. I can see a sailboat and a kayaker. It's really pretty.
Outside Landscape	4	There is so much to see looking out the window back towards the city. From watching people having their lunch down below outside the restaurant, people using the exhibits on the Pier, and all the way at the end of the Pier the Embarcadero and all the traffic. I just saw the historic F train go by. Up the hill behind it you can see where there was a quarry in the 1800s and the early 1900s. If you look up at the buildings you'll see a bunch that are right above an exposed…

Participants

An assistant evaluator stood at the entrance of the Observatory and approached every third visitor, who appeared 8 years-old or older, as s/he was about to enter. The evaluator asked that visitor if s/he would be interested in giving feedback on a mobile app the Exploratorium was developing. If the visitor was with another person, the evaluator welcomed both visitors to participate. If the visitor was a minor, the recruiter first asked the accompanying adult for permission to talk with the youth. Visitors willing to participate were asked to sign a consent form allowing us to capture and analyze any recordings they made with Open Conversation.

Halfway through the formative evaluation, we realized that most of the participants were male, at which point we changed our recruitment to select and approach the female in the third group who entered the gallery. If there were no females in the selected group, the recruiter would ask a male visitor. However, even with this change, we were unable to obtain a balanced sample. This is mainly because many of the females who came either refused participation or came with very young children, who, we felt, would prevent them from engaging with the app. Likewise, it was difficult to recruit minors because it was oftentimes challenging to identify or find an adult who could consent to their participation. The demographics of our study participants are shown in Table 2. Note that in all but two cases, a single person used the app. Even when another person consented to trying Open Conversation, s/he would very early on disengage. The only two exceptions were a dyad of teenage girls and an adult peer group who used the app together throughout their gallery visits.

Table 2. Demographics of study participants

Age Group	Count
emerging adult (18-20s)	8[++]
adult	13
senior	2
child	2
teenager	2[+]
Total	27

Gender	Count
Female	11[+]
Male	16[++]
Total	27

[+] The pair of teenage girls was counted only once, as one study participant.
[++] The two male emerging adults who used the app together were considered as one group and counted once.

Data Collected

In total, we recruited 27 visitors over nine days in April and June 2015 to participate in this formative evaluation and collected the following data for each participant:

- Timing and tracking data. The assistant evaluator noted when the visitor stopped and left each exhibit during his/her gallery visit. This information was collected for every participant except the first recruit, when we were still refining the timing and tracking protocol.
- IPS location data. A logging application installed on the smartphone recorded the IPS-derived location information for all but one of the 27 participants[1]. These logs were later downloaded and its latitude / longitude data matched with the data collected from Roundware to check location information.
- Roundware data. Roundware logged every time an audio file played and when a visitor made an audio recording.
- Interview data. The assistant evaluator interviewed the study participant(s) immediately after s/he indicated that s/he had finished using Open Conversation. Out of the 27 participants, we interviewed 26 visitors. (We chose to end one of the interviews early because technical problems with the application severely compromised that visitor's experience.)

RESULTS AND DISCUSSION

What did visitors do with the app?

Only one of the 27 study participants chose to use earbuds, which we offered to all of our visitors at the start of their experience. Everyone held the mobile phone in his/her hands throughout the gallery visit.

Listening to comments others left

All of the 26 visitors whom we interviewed[2] listened to the comments and stories Open Conversation played as they visited the Observatory gallery. On average, visitors listened to 19 (mean) or 15.5 (median) recordings, and Roundware audio was playing during 65% (mean), or 58% (median), of their gallery visit. (See Figure 4 and Figure 5 for the respective histogram.) A closer look at the Roundware data further revealed that once the IPS-Roundware system acquired location and started streaming audio,

[1] The logging app was not available for the first study participant.
[2] Twenty-seven visitors tried Open Conversation, but one participant group reported that the app only played music throughout the entire visit; we did not interview this group.

the audio comments played one after the other with only a very short period of silence between. Figure 6 shows the sequence of audio played to visitors. Although there was a pause button, few visitors used it, and data suggest that visitors may have simply forgotten that there was one: one visitor wished there were a way to stop the audio, and another visitor only remembered the pause button during her interview.

Figure 4. Histogram of the number of audio recordings played.

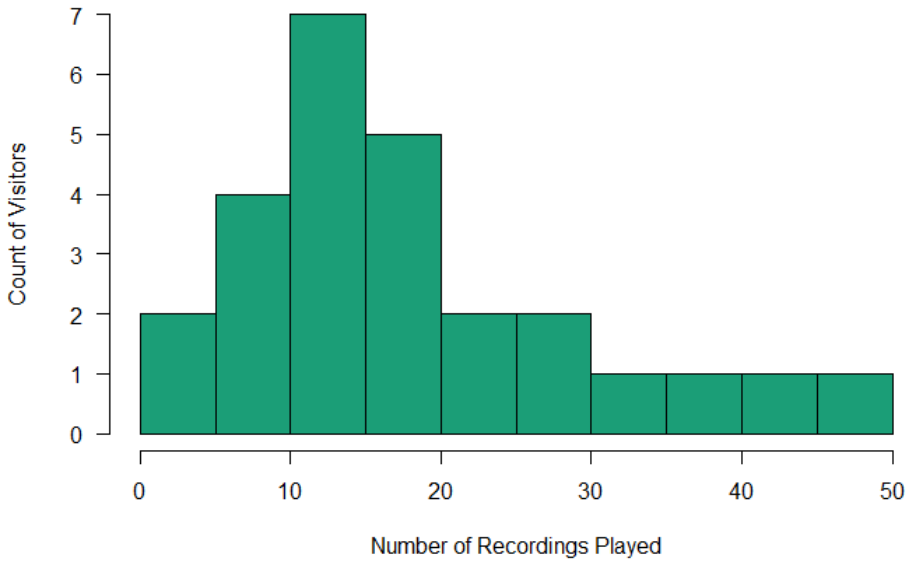


Figure 5. Histogram of the fraction of the time an Open Conversation audio was playing.

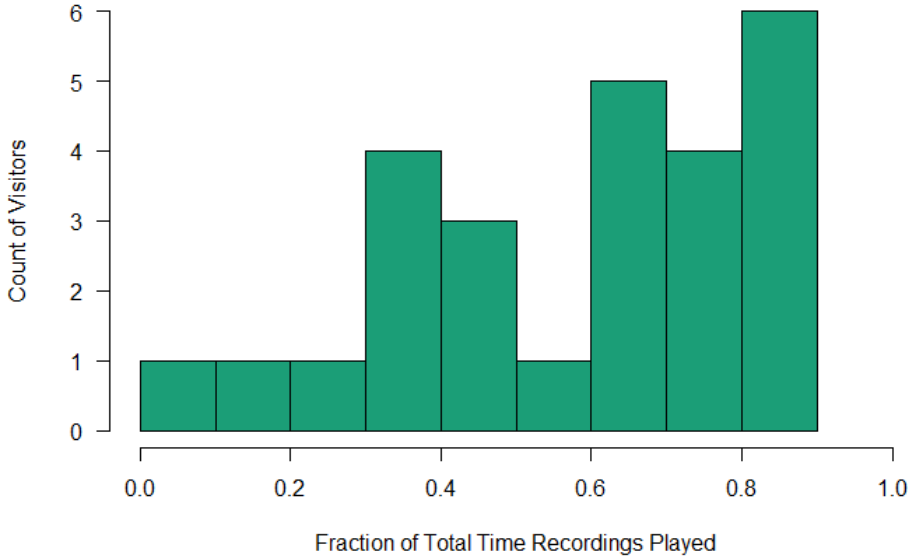


Figure 6. Sequence of audio recordings played

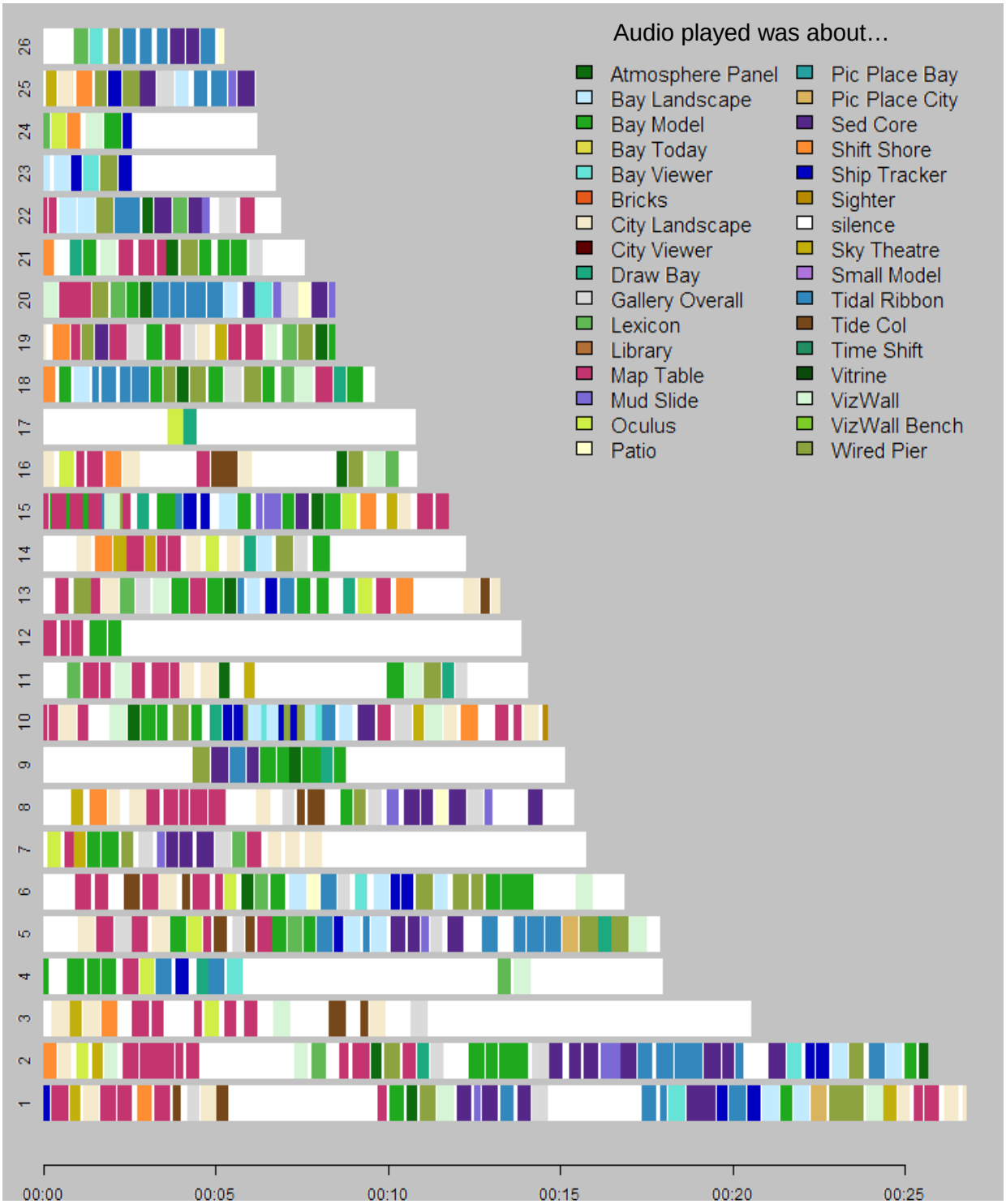


Table 3 tallies the types of comments that Open Conversation played. Because the Roundware database was seeded with many more stories and comments about particular exhibits, not surprisingly, most of the audio visitors heard were about exhibits.

Table 3. Types of comments and stories played to a visitor

Comments about…	Percentage of the comments played
An exhibit	84%
The gallery in general	4%
The landscape outside	12%

Comment from…	Percentage of the comments played
Staff	49%
Visitor	30%
Scientist	10%
Artist	10%

Recording their own comments

Alternatively, only a small minority (5 out of 26) of the study participants recorded their own comments on IPS-Roundware. All but one of these visitors left one recording; this one person made three during her gallery visit. Of the seven total comments the study participants recorded, five were of a particular exhibit the visitor used, one was about the outside view, and one, "This is so cool," could not be connected to any specific aspect of their experience.

The large majority (21 out of 26) of the visitors, however, did not make any recording. The most common reason (7 out of 21) visitors gave for not leaving a recording was that they felt they had nothing to add. Other types of responses are listed in Table 4.

Table 4. Reasons for not recording a comment

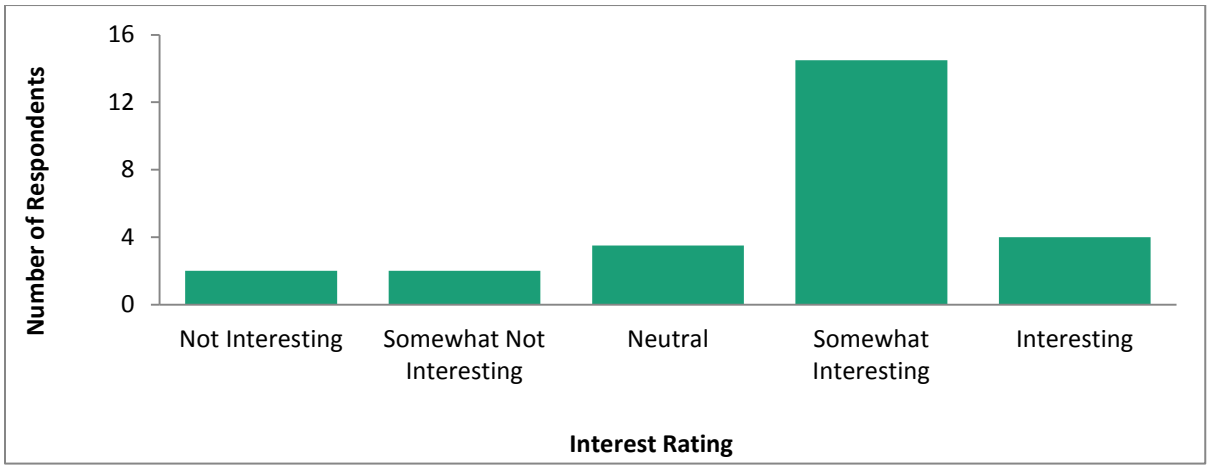
Reason Given	Count (out of 21 visitor who did not record a comment)
I had nothing to add	7
It never occurred to me	4
There were technical difficulties	3
I was too shy	2
I didn't know how	1
It was too much work	1

The above data all suggest that Open Conversation was predominantly a one-way conversation about exhibits, with the visitor listening to what others said.

What visitors found worthwhile

Overall, visitors (N = 26) rated their Open Conversation experience to be *Somewhat Interesting*, or 4 (median) and 3.6 (mean) out of a 5 point Likert scale with 5 being *Interesting.* (See Figure 7.)

Figure 7. Histogram of visitors' self-reported interest ratings



What visitors found worthwhile (self-reports)

When asked, visitors gave different reasons for finding the experience worthwhile. The most frequent explanations included appreciating hearing other viewpoints, a chance to listen rather than to read, and getting additional information. Examples of visitors' responses are provided in Table 5. Other, less frequent reasons visitors gave were: They liked a particular feature of the interface such as the ability to select comments left by scientists (5 respondents), the casual tone of the stories (4 respondents), and the social interactions the app afforded (4 respondents).

Comparing gallery visits with and without Open Conversation

To gauge if Open Conversation promoted visitor engagement with the Observatory gallery, we compared the timing and tracking data from this study's participants to data collected from a separate baseline study, in which visitors were asked to wear a mobile phone and tracked by an evaluator as they visited the Observatory. In particular, we looked at their dwell time in the gallery[3] and the number of exhibits they stopped at.

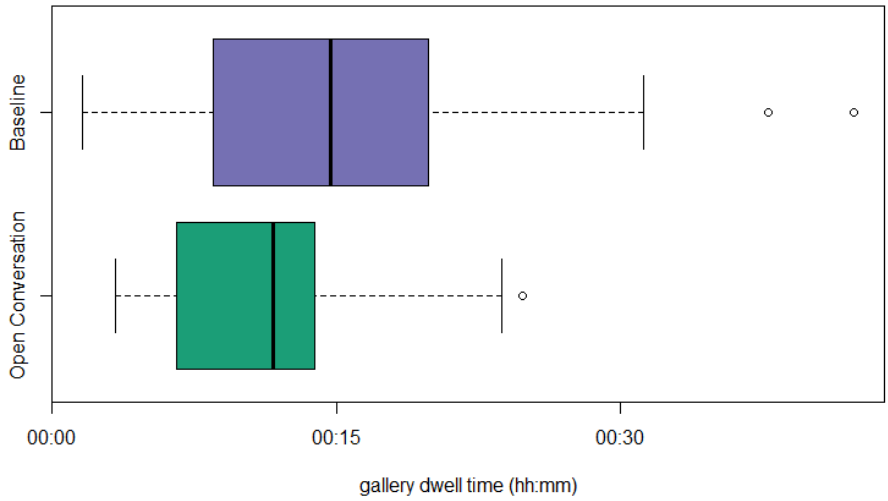
Gallery Dwell Time. On average, Open Conversation visitors spent 0:11:39 (median) in the Observatory, while the baseline group was in the gallery for 0:14:42 (median). Figure 8 shows the boxplot for each group. A Mann-Whitney test found no statistical difference between the two groups; U = 246, p > .05.

[3] To adjust for the period at the start and end of the visit, when the study participant may still be talking with the recruiter or learning the app, we defined the gallery dwell time to be the time from the visitor's first exhibit stop to his/her last exhibit stop.

Table 5. What visitors found worthwhile with example quotes from visitors' interview responses.

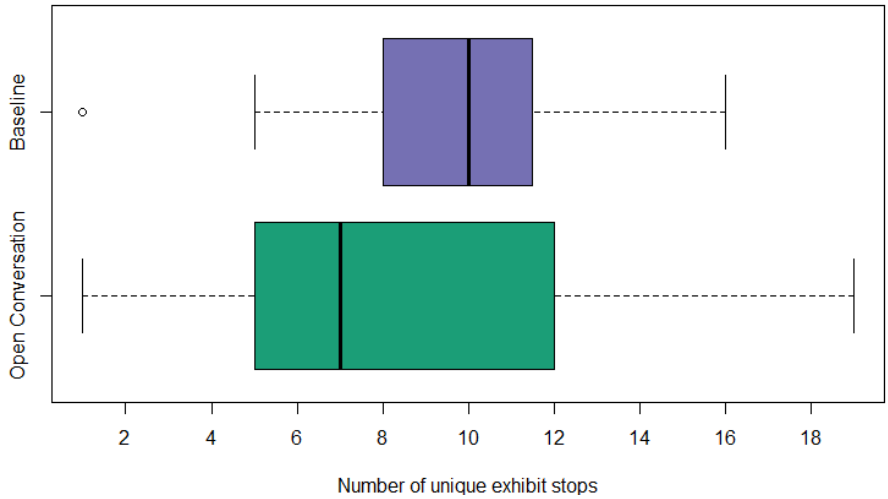
Visitors appreciated hearing another person's perspective (13 / 26 respondents)
Visitor2: Interesting way that someone else's experience. For example, if someone described dolphins and whales they saw out the window, that experience could be shared because they saw it
Visitor8: There is a lot of hands-on stuff to do. It would be cool to see what other people thought of it.
Visitor15: When I find places/things interesting it would be good to hear different points of view about it.
Visitor18: I enjoyed it. It was interesting to hear different perspectives.
Visitor26: you could see in other people's eyes, see what they saw
Visitor27: When I was listening I would go over to what they were talking about and when they talked about their experience I would think about why they thought that and I would try to think about it the way they did.
Visitor28: I liked the artists and the scientists too because science is my favorite subject also I like art so I can relate. Also I liked how they saw things differently but at the same time in the same way.
The recordings helped them notice what they would have overlooked otherwise (10 / 26 respondents)
Visitor0.5: [I would have] easily overlooked those things with friends
Visitor1: At the tidal gates (tidal ribbon) I overlooked something and it said something I would have otherwise missed. At [core sediments] I was trying to understand other's experience, It helps me more than what I would have normally done. Makes exploring easier.
Visitor3: It showed me things I wouldn't have been looking for.
Visitor9: The ones pointing out things I wouldn't have seen on my own. Like the one about the map table and the secret history map. I searched to find that map and I wouldn't even have known about it.
Visitor23: I wouldn't have noticed like the maps. Same with the column (tidal column)… Some of the things mentioned I wouldn't have paid attention to.
There was something appealing about listening versus reading (9 / 26 respondents)
Visitor3: I am an audio learning so it helps me when people explain it and say this is what's going on.
Visitor6: You don't need to stand and read, and you can use your hands on the exhibits more. That I like.
Visitor13: It was more realistic than reading. When you first come in [to the museum] you want to read everything, but here [at the end of your trip], you get information overload.
Visitor22: I'm a little bit lazy. My husband reads every little thing. I like the idea of someone telling me.
Study participants liked the additional information the comments provided (8 / 26 respondents)
Visitor6: The ones with new information/knowledge. They weren't too long.
Visitor15: Someone who already knows the information is giving me the information I would have had to search for otherwise. So the information felt more readily available. The more exposed you are to information the more you learn.
Visitor18: It's information I want to hear about.
Visitor24: One thing I liked that I didn't expect was when I got more extensions of information like what goes on in the bay… When I was by the windows it was talking about how the Chinese shipped their dead bodies' home and how they did that. It was interesting because it was a bit of history not contained within this space, like a broader sense of information that I wasn't expecting to learn.

Figure 8. Boxplot of gallery dwell time for visitors using Open Conversation versus visitors who simply carried a mobile device.



Number of Exhibits Visited. When we compared the number of exhibits visitors stopped at, we found no statistically significant difference between the Open Conversation users (M = 7.88, SD = 4.93) and the visitors who were simply asked to wear a mobile phone (M = 9.37, SD = 3.31); $t(43.5) = -1.28$, $p > .05$. See Figure 9 for the boxplot.

Figure 9. Boxplot of the number of unique exhibits visitors stopped at for visitors using and not using Open Conversation.



Our analyses of visitors' timing and tracking data, therefore, did not provide any evidence that the IPS-Roundware app led to increased exhibit use and engagement in the Observatory. Although not a statistically significant result, we note that Open Conversation visitors actually spent less time in the gallery and used fewer exhibits. We suspect that usability issues may have been a major source of

frustration and confusion that compromised visitors' experience with Open Conversation and hence with their overall engagement in the Observatory.

Points of frustration and confusion

When we examined visitors' interview responses for points of frustration and confusion, we found a set of usability issues, which are listed in Table 6.

Table 6. Usability issues with Open Conversation with example quotes from visitors' interviews.

The audio that played did not correspond with the exhibit visitors were using or what visitors were looking at (18 / 26 respondents)
Visitor2: I feel like it may or may not be relevant to what I was looking at or what I was listening to at any given moment
Visitor3: My only concern is that it wasn't about the place I was at. Tidal Ribbon was being talked about when I was at the Map Table. I didn't understand what the comment was about because I hadn't seen it yet.
Visitor5: It was hard to figure out what was going on. I realized that it was about specific locations. I found a couple after [I heard the exhibits, but not while they were discussed.]
Visitor9: There was a disconnect between what I was seeing and what was being talked about… When I was at the map table it was talking about the core and then when I went to the core it started talking about something else.
Visitor11: When I was in the sky theater it started talking about the view. I felt conflicted, should I leave and follow the comments? So then I left and then it started talking about sky theater, which is what I wanted to hear about from the start.
Visitor17: I'm not even sure what device they are using…Like what exhibit they are using and talking about.
Visitor18: It wasn't obvious to me that it knew where I was so it was disorienting. I don't know what you're looking at or what you're talking about. (the person speaking on the app) It just seemed like it was more effort than it was worth.
Visitor23: if it matched more where I was at. Otherwise it's too confusing.
Visitor24: It was confusing because I didn't know what they were talking about or what they were looking at
The app did not support a function visitors wanted or needed (15 / 26 respondents)
Visitor0.5: Wish it was a next button for the next exhibit because it was jumping. A pause or button for the next exhibit instead of the audio going on and on
Visitor6: If I knew how much information you have [for each exhibit] then I can estimate if I want to wait [and listen] or not.
Visitor9: If you want to know exactly where the exhibits are there isn't any way to figure that out from the app.
Visitor13: It was confusing when there were different voices talking…I wanted it to just be one person or voice.
Visitor21: I didn't like the music in the background I felt like it was distracting from the speaking and disorienting.
Visitor22: I hoped it would go to the next story soon. I would have liked to be able to hear it again but it doesn't have the choice to go back.
Visitor23: Also, if you could wear it around your neck instead of carrying it. At one point I wanted to take a picture but I was holding that thing (the device).

Visitor26: maybe it would be a good idea to have a fast forward button or a go back button
There were additional problems beyond geolocation issues (12 / 26 respondents)
Visitor6: Loading time was longer than I expected.
Visitor7: It crashed continuously.
Visitor11: When I was trying to comment the screen went black and it didn't respond to my touch so I couldn't pause or record but the music kept playing
Visitor23: I couldn't hear them all very well. I was so concerned about not being noisy. I should have taken the headphones (which were offered to all the study participants).
Visitor27: It was hard to understand them [the audio] at times.
Visitor29: then it stopped working.

In addition to these usability issues, we found that some visitors did not appreciate the audio stories' content and found that Open Conversation interfered with their preferred gallery experience. A count and examples of these visitors' comments are listed in Table 7.

Table 7. Other factors that visitors found frustrating with example quotes.

The audio stories' content was not interesting (11 / 26 respondents)
Visitor5: [I wanted a] more technical explanation. What I think I heard was more a reaction to what people were observing.
Visitor6: Non-visitors gave more background, which was nice… What the other visitors said wasn't really interesting.
Visitor11: I couldn't relate to most comments. Some of the comments were helpful but it forces other people's perception on you.
Visitor12: I didn't enjoy listening to the comments … I even tried to select (pressed the modify button) only artists and staff but it still seemed random. I am more interested in the artists and staff comments. I wanted more background and professional opinions and comments. Like the goal of the exhibits what they are trying to show, the ideas behind the exhibits and how they came up with them.
Visitor17: As I understand it you just press play to hear the voices. Why? The comments from the general [populous] were not interesting… If I want to interact more with the exhibits, the plaques should be enough. It should have all you need to know.
Visitor26: One guy was talking about the shadows on the wall which I didn't understand.
Open Conversation disrupted their gallery experience (8 / 26 respondents)
Visitor10: I like just figuring things out on my own and going at my own pace.
Visitor16: I couldn't hear with my mom and the thing talking at the same time.
Visitor21: It was distracting from the actual exhibit.
Visitor23: At one point I wanted to take a picture but I was holding that thing (the device).

How did mismatches affect the visitor experience?

The dominant usability issue seemed to lie with the lack of exhibit level resolution the prototype IPS was able to provide. More than two-thirds of the visitors we interviewed complained about the geolocation accuracy. To better understand how the lack of exhibit level resolution affected the visitor experience, we took a closer look at visitors' reactions to mismatches between the audio played and the exhibit they were using.

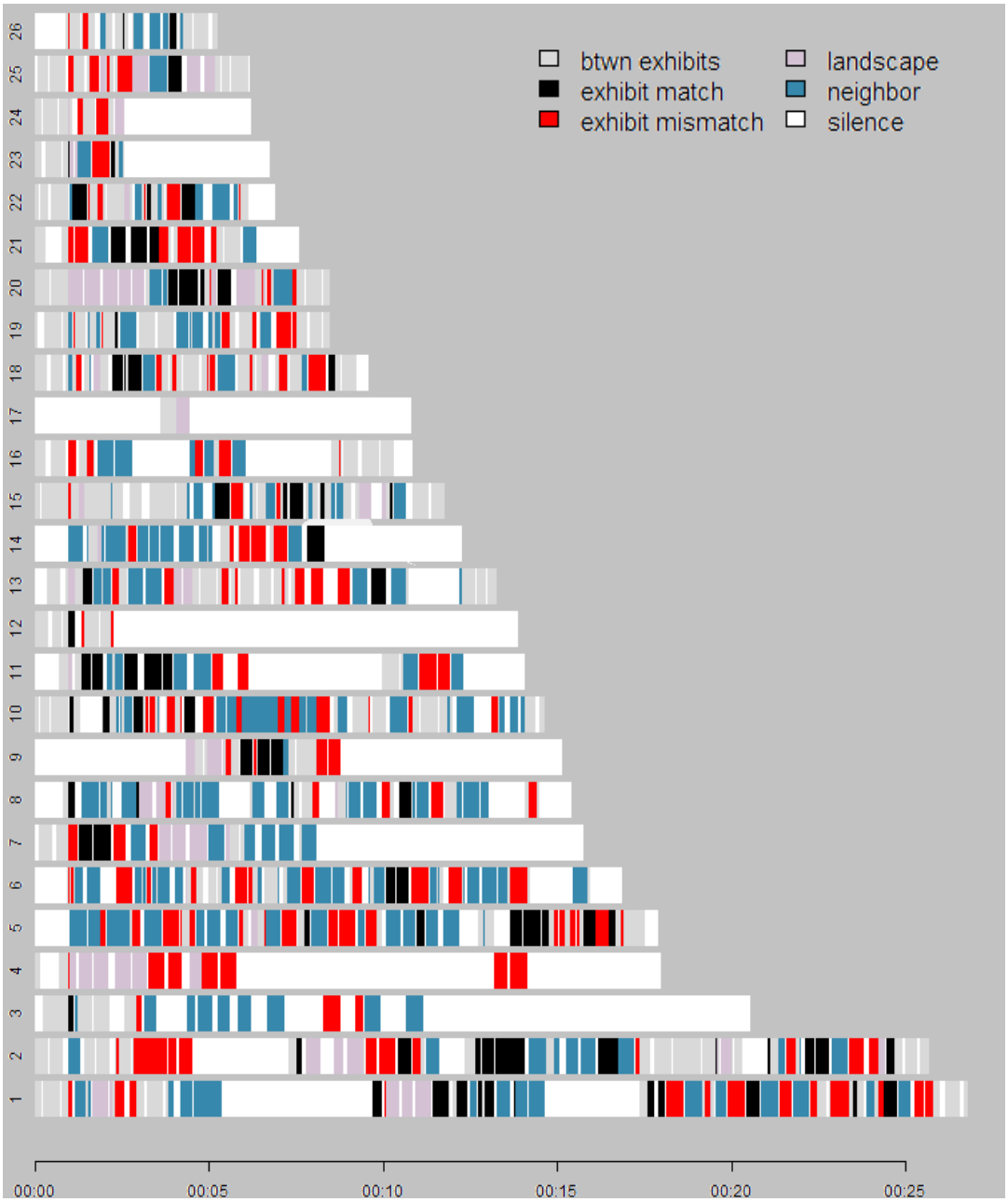
First, we looked in the interview transcripts at visitors' self-reported responses for the range of reactions to a misalignment. Study participants talked about:

- doing nothing,
- trying unsuccessfully to force the app to play an audio clip about what the exhibit they were using, and
- looking around for the exhibit the comment or story was describing.

Second, we examined the Roundware logs and the tracking and timing data to see if the audio played aligned with the exhibit visitors were at. We found, when a visitor using Open Conversation was at an exhibit or looking at the outside landscape and an audio recording was played, that audio matched the exhibit s/he was using 20% of the time, and described an adjacent exhibit or the visible landscape 43% of the time. Overall, the audio that was playing pertained to something nearby or in view on average 63% (mean), or 67% (median), of the time when a visitor was using an exhibit. But, approximately one-third of the time, the audio that played had nothing to do with the exhibit or the scene visible to visitors.

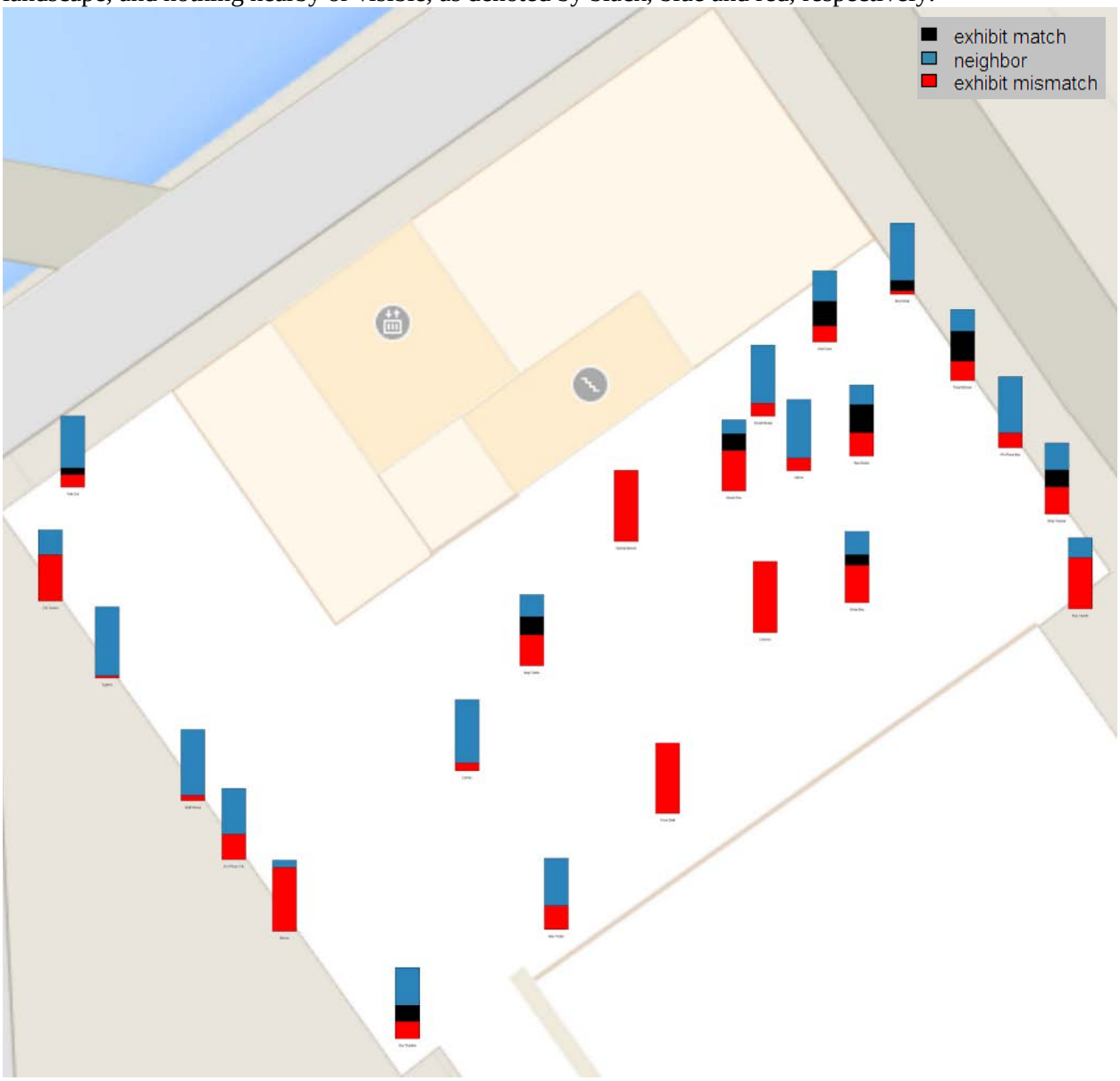
Figure 10 shows the discrepancies between the exhibit location of the visitor, as observed by the human data collector, and the audio recording the visitor heard while there. Note that the misalignments, denoted with red bars, are distributed throughout the visitor experience instead of clustering at a certain time, for example, at the end of anyone's visit.

Figure 10. The discrepancy between the exhibit visitors used and the audio heard. Black bars denote when Open Conversation played an audio recording about the exhibit the visitor was using, while the blue bars indicate when the audio recording played described an adjacent exhibit, the landscape visible to the visitor, or the general gallery. Red signifies a misalignment; that is, the audio played talked about something that was not nearby or visible from where the visitor was standing. Lastly, light grey indicates when a visitor was in transit between exhibits and not looking at the landscape, white denotes when no audio was played, and light purple shows when a visitor was looking outside but the audio played was about an exhibit; these times were not included in the calculation of agreement.



Plotting the discrepancies according to exhibit location (Figure 11), we found that there were areas with better audio alignment than others. Improving resolution may help ameliorate the amount of misalignment a visitor experiences.

Figure 11. The percentage of agreement between the exhibit visitors used and the audio heard for each exhibit according to its physical location in the Observatory. Each stacked bar graph indicates the percentage of time when the audio that played was about the exhibit, a nearby exhibit or visible landscape, and nothing nearby or visible, as denoted by black, blue and red, respectively.



However, despite a concentrated effort to attain exhibit-level resolution with our prototype Wi-Fi based indoor positioning system, we were unable to achieve position data better than a few meters. And, exhibit-level resolution may remain years out (Johnson, Adams Becker, & Freeman, 2013). In lieu of an immediate technical solution, there may be other ways to help align what is played with what visitors see. These all involve asking and helping visitors themselves to 'bridge' the final few meters that an indoor positioning system itself cannot resolve and locate the exact exhibit themselves. Possible solutions along this line may include redesigning Open Conversation and Roundware to provide:

- a photo of the exhibit the audio is describing,
- a map to help visitors locate different exhibits,
- time between stories (or a more prominent pause button) to allow visitors to locate different exhibits, and/or
- a 'next' button so visitors can align the audio with what is in front of them.

These workaround may help address some of the usability issues that arise from the current lack of exhibit level geolocation for an IPS-Roundware system.

SUMMARY

This formative evaluation on the first prototype of the IPS-Roundware application, Open Conversation, found:

- Open Conversation provided predominantly a listening experience. All of the visitors we interviewed used Open Conversation to listen to other people's audio recordings, while few (5 out of 26 participants) left any recordings of their own. Furthermore, there were short periods of silence between stories once the system connected to the server and began streaming audio, which may have discouraged visitors from 'interrupting' to contribute to the conversation.
- In large part, Open Conversation was about exhibits, as opposed to the view outside or the qualities of the overall gallery. This is reflected in both the type of stories and comments visitors heard and recorded, which largely pertained to particular exhibits. This was not too surprising since many of the comments collected from staff, scientists, artists, and a few visitors, before the formative evaluation to seed the application were about exhibits. Because we did not have a balanced number of stories related to the outdoor landscape or gallery nor a means for the visitor to choose the type they wanted to listen to, this evaluation could not determine if visitors had a preference for one type of story or another, whether it be about exhibits, the overall gallery, or the view of the landscape outside.
- Visitors found their Open Conversation experience worthwhile for several reasons. The three most frequently mentioned were: 1) They appreciated hearing another person's perspective; 2) there was something appealing about listening versus reading; 3) They liked the additional information the comments provided.
- Some of the visitor responses indicated that the app may have encouraged visitors to spend more time exploring the gallery. These interview comments echo the findings from an evaluation study conducted on a Roundware app developed and used at the outdoor sculpture garden at the Museum of Fine Arts in San Francisco (Girardeau, Beaman, Pressley, & Reinier, 2015). However, when we compared the gallery dwell time and the number of exhibits visited for our study participants, who used Open Conversation, with gallery visitors recruited for a baseline case, who did not use Open

Conversation, we found no statistical difference. That is, we could not find evidence that the app encouraged visitors to stay longer or visit more exhibits in the Observatory.

- This study surfaced a number of issues that detracted visitors from fully engaging with the app and the gallery. Participants reported feeling frustrated and/or confused by 1) misalignments between the exhibit they were at and the comments they heard, 2) a feature (or lack thereof) of Open Conversation, 3) other issues of the IPS-Roundware system (e.g., a lag in playback time) that were not related to the indoor positioning, 4) uninteresting comments others left, and 5) the nature of the interaction the application afforded.
- This formative evaluation found that over two-thirds of the (26) visitors we interviewed reported feeling frustrated and/or confused by the mismatches between what they heard and their exhibit experience. The IPS did provide cluster-level resolution for parts of the Observatory (i.e., the recording played corresponded with the exhibit the visitor was at, an adjacent exhibit, or the view visible from that exhibit). However, exhibit level resolution (i.e. the recording played matched the exhibit the visitor was at) was difficult to achieve for most areas within the gallery. Going into this study, we were not sure how important exhibit level resolution was to this type of visitor experience. These results suggest that resolution better than 2 meters is an important criterion for Open Conversation for galleries such as the Observatory.

ACKNOWLEDGEMENTS

The authors would like to thank Katherine Nammacher and Stephanie Stokes for recruiting visitors and for administering the interviews for this study. The custom-built IPS-Roundware app, Open Conversation, was developed by Halsey Burgund and Claire Pillsbury. Douglas Thistlewolf configured the IPS with the generous help and technical expertise of engineers from our technical partner, Qualcomm.

This material is based upon work supported by the National Science Foundation under Grant number 1346664. 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 National Science Foundation.

REFERENCES

Girardeau, C., Beaman, A., Pressley, S., & Reinier, J. (2015). Voices : FAMSF : Testing a new model of interpretive technology at the Fine Arts Museums of San Francisco. *MW2015 : Museums and the Web 2015*.

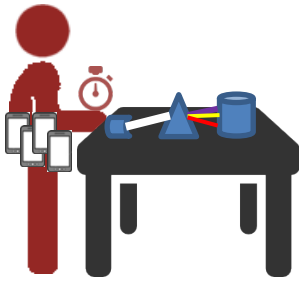
Johnson, L., Adams Becker, S., & Freeman, A. (2013). *The NMC Horizon Report: 2013 Museum Edition*. Austin, Texas, USA: The New Media Consortium.

APPENDIX A: Ground Truth Protocol

The following outlines the procedures for creating an exhibit map using an indoor positioning system. The resulting location, in latitude and longitude, is as seen by the IPS and does not always reflect the exhibit's physical location.

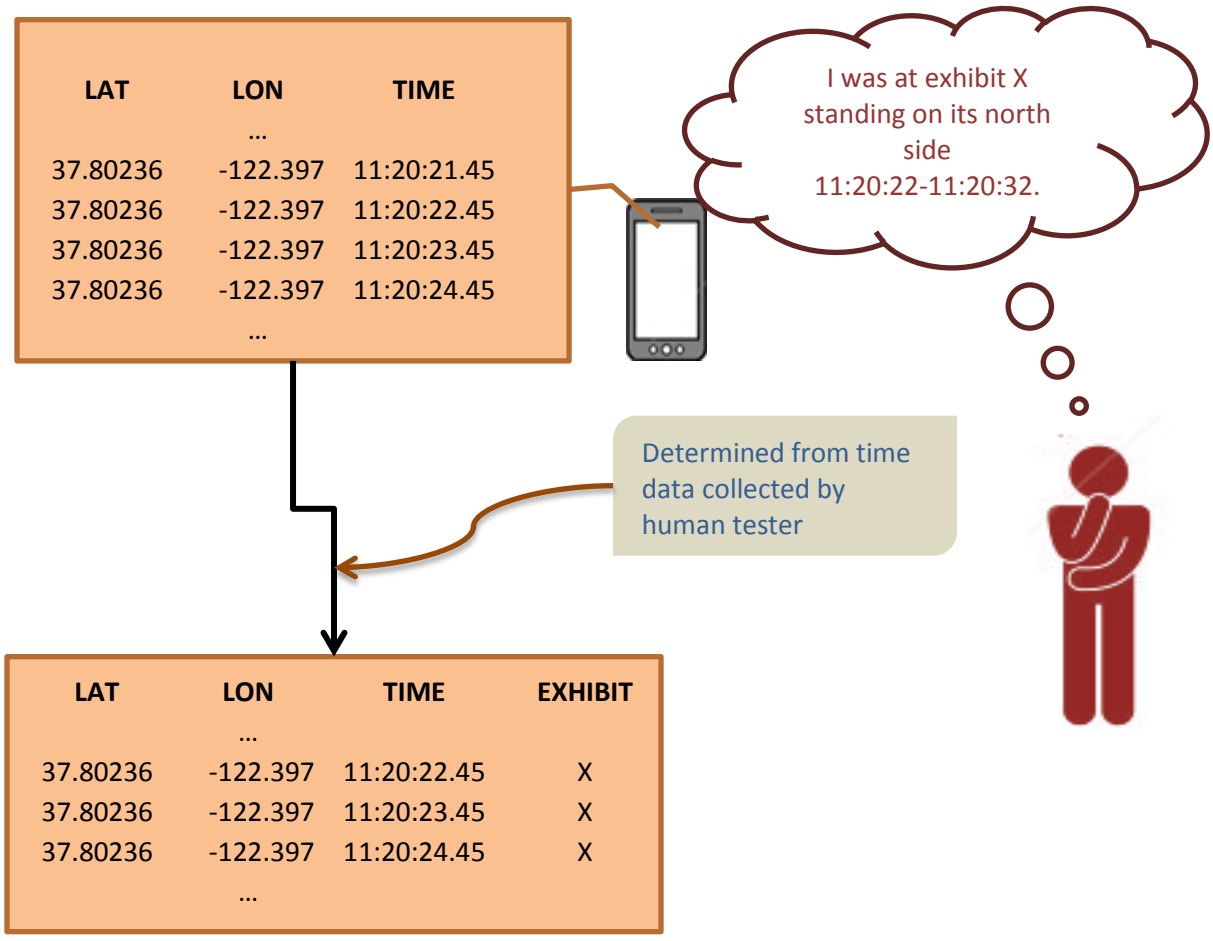
Method

1. Survey the gallery
    a. identify every exhibit
    b. identify the use positions for each exhibit. For example, if the exhibit is a four-sided tabletop in the middle of the gallery, there will be four use positions, one on each side.
2. Turn on the IPS logging app and note when the app was launched (to the second) for each IPS-enabled tracking device. The tracking log should record the latitude, longitude of where the device is, according to the IPS, and the current time (to the millisecond).
3. Wear the tracking devices equally spaced around your waist.
4. For each exhibit's use position
    a. Stop and note when you've stepped into that position.
    b. Use the exhibit from that position for 30 seconds.
    c. Step away from the exhibit's use position and note when you've done so.



Data Processing

1. Download the IPS data logs from each device.
2. Synchronize the data log from each device with the timing notes you took to determine when that device was at each exhibit's use position.
3. Using the data logged during each exhibit stop including all use positions for all the devices, calculate the centroid (i.e., the mean latitude and the mean longitude). This is the position of each exhibit according to the IPS.


LAT
LON
TIME
...
37.80236
-122.397
11:20:21.45
37.80236
-122.397
11:20:22.45
37.80236
-122.397
11:20:23.45
37.80236
-122.397
11:20:24.45
...
I was at exhibit X standing on its north side 11:20:22-11:20:32.
Determined from time data collected by human tester
LAT
LON
TIME
EXHIBIT
...
37.80236
-122.397
11:20:22.45
X
37.80236
-122.397
11:20:23.45
X
37.80236
-122.397
11:20:24.45
X
...