

Group-Based Museum Audio Dramas for Well-Being

Charles Callaway^{1,*} and Oliviero Stock²

¹TrentoRISE, via Sommarive, 20, Povo (TN) 38100, Italy

²FBK-irst, via Sommarive, 18, Povo (TN) 38100, Italy

Abstract

Well-being in a small group can be tied to how much its members interact. Small group tours are social occasions, and the discussion that ensues has been shown by ethnographers to be important for a more enriching experience. Increasing conversation can thus be seen as a way to improve social and psychological well-being. We present DramaTric, a mobile presentation system that delivers hour-long dramas to groups in museums. DramaTric gets sensor data from its environment and analyzes group behavior to deliver dynamically adapted dramatic scenes designed to stimulate conversation. Each scene contains slight differences in the story, leading visitors to understand their own drama only by talking with other group members. We describe an experiment with a full-scale drama to test if switching from presenting a drama with one technique to another results in more conversation. This shows that by using adaptive techniques we can modify social behavior, which can in turn promote well-being.

Received on 03 December 2013; accepted on 03 February 2014; published on 04 March 2014

Keywords: Groups, cultural heritage, mobile and ubiquitous systems, ambient sensors and systems, drama.

Copyright © 2014 C. Callaway and O. Stock, licensed to ICST. This is an open access article distributed under the terms of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi:10.4108/amsys.1.3.e6

1. Introduction

Improvements in the well-being of social groups can be gauged socially, emotionally and psychologically. But for implemented real-time mobile systems, such improvements must be measured quantitatively in order to have an impact on the system's adaptive responses to ongoing group behavior. In cultural heritage settings, such as museums and art galleries, social groups are accustomed to interacting via conversation. In fact, social groups are the dominant means of museum attendance [10], and conversation has been shown to be important in whether groups in such settings experience more positive visits [9]. We thus see adaptations by the system as a purposeful way to increase conversation, and thus the well-being of groups touring cultural heritage sites.

Additionally, the deployment of audio guides in the last two decades has been characterized by fact-based exhibit descriptions and the semi-isolation of group members whose audio guides are not aware of each other. Such implementations can lead to a decrease in engagement on the part of group members which is an

important factor in how much conversation takes place. A different approach to increasing engagement and thus conversation is to use presentations that evoke emotion, such as performed narrative like drama or film [2].

We created the DRAMA^{TRIC} system (Drama Tension Release by Inducing Conversation) which uses a novel group-oriented approach integrating sensors and analysis of the resulting sensor data into a coordinated narrative system for mobile devices (specifically, smartphones and tablets as presentation devices) in a museum instrumented for ambient intelligence. DRAMA^{TRIC} is based on adaptive narration, where by adaptivity we mean using observed and inferred characteristics of group behavior to choose when and which drama-based presentation to show, as well as how to link successive dramatic scenes together. DRAMA^{TRIC} allows a group of visitors to move freely around a museum along any path they choose, and produces a larger drama by stitching together smaller drama segments and ensuring a coherent narrative regardless of that path. Sensors both in the museum and worn by the visitors allow DRAMA^{TRIC} to know (a) where individual group members are, (b) objective characteristics such as proximity between members, speaker, and length of conversational turns, and (c)

*Corresponding author. Email: ccallawa@gmail.com

inferred characteristics, such as who is the group leader, or how much time they spend near each other.

The specific method we use to increase conversation involves presenting slightly varied narratives for different individuals or subgroups of a small group while they are listening to the same dramatic scene. We purposefully create these variations by allowing some visitors to hear parts of the narration that others do not. We thus create conditions where visitors can only understand each drama segment by filling in the details that their fellows lack, resulting in subsequent conversation. Below we describe this methodology in further detail, followed by a description of the hardware, architecture, and adaptivity mechanism, and the description of an experiment in an actual museum that shows narrative variations can lead to significant differences in the amount of conversation immediately after the conclusion of each audio drama. We thus show that it is possible to increase conversation within a group, lending support to the hypothesis for increasing group well-being.

2. Narrative Variations and Drama Presentation

The methodology we have created for promoting interaction, and especially conversation, is to give group members dramatic presentations with a particular type of unresolved narrative tension that forces them to talk to each other to relieve that tension. We use the term tension to indicate that the visitor listening to the audio drama is engaged in the museum experience and hears a drama where certain information critical to understanding the narrative is removed in a *coordinated way*, or that it is different from what other group members hear. Group members will then notice that this content in our dramatic narrative variations is incomplete [3], and then act on this by questioning or discussing with their fellow group members, actively leading to a more general conversation ranging from the themes showcased in the museum to those themes in their daily lives. In other words, we want to purposefully increase narrative tension in the audio dramas heard by members of a small group, and let the natural curiosity inspired by that narrative tension spur the resolution of that tension, *i.e.*, by finding the missing narrative information that completes the story.

As our specific mechanism, we present slightly different audio dramas to different group members by selectively withholding information from some members of a group but not others. Thus person A hears something important that person B doesn't, and vice versa. We connect narrative variations to increased conversation by purposefully writing the dramas and manipulating the audio in such a way that each scene is missing one of the key points that narratively completes that scene. By carefully structuring these "information

gaps", each person can be given content that their fellow group members require in order for them to completely understand the events in their own scene. We have developed three specific techniques to provide for this narrative content variation:

- *Telephone*: A one-sided conversation style where the audio of only one character can be heard as if talking via telephone. Pauses, music or sound effects are inserted when the other character, whom we can't hear, is speaking.
- *Audio blurring*: This technique "overlays" the dialogue at selected points with some source of ambient noise (*e.g.*, seagulls screaming, children yelling, the sound of waves or wind, etc.). The dialogue at these key points is thus rendered unintelligible. However, the group members can still tell that the characters are conversing, as the volume of the interference is just below the volume of the dialogue.
- *Point of View Change*: When there is a social conflict between two or more characters, we can allow each character to present their own viewpoint without interference. We thus have two or more monologues instead of a dialogue that reflects the point of view of specific characters, while other group members hear the point of view of a different character.

Museums typically consist of a large number of exhibits organized thematically in a number of different exhibit halls. Visitors can choose which route to take between the halls, skipping some exhibits entirely and exploring the others in whatever order they prefer. A principal difficulty is thus to maintain coherence regardless of which order the drama segments are heard, and even if some dramas are completely skipped. In our approach, we present a series of short, self-contained dramas for visitors in front of certain major exhibits or important locations, where the specific technique of narrative variation is based on adaptivity criteria described below. We found that when our drama tour was fully completed, an overall drama across 5 exhibit halls and consisting of 14 individual short audio dramas required between 35 and 55 minutes.

We utilized 4 distinct types of audio drama, each with its own role in maintaining the story arc and tension:

- Fixed drama segments: 30-second to 2-minute-long self-contained dramas that are standardized for every visitor. These function to set up the drama (introduction and initial development) and to bring the drama to a conclusion.
- Primary drama segments: 1- to 2-minute-long self-contained dramas with a cast of characters and an identifiable plot, written for a particular exhibit. Immediately after the completion of

a primary drama segment, there is a 1-minute observation period of the behavior of the group members, and the results of that observation are used to adaptively select a set of linking segments and a technique (described below) to employ for the next primary drama segment.

- Secondary drama segments: 1-minute long dramas without narrative variations and not followed by an observation period. They serve to introduce the visitors to a new exhibit hall, artifacts of lesser importance, or to new characters who will play a role in later dramatic presentations.
- Linking segments: Because each drama segment is independent and the order they are played in is determined by the path visitors take through the museum, the end of one dramatic presentations may not match up coherently to the start of the second without some means of tying the scenes together. We thus use multiple *linking* segments, lasting about 10 seconds each, between the drama segments to provide continuity and to help the visitors perceive that the system is reacting adaptively to their behavior.

Because high quality acting and production values are critical to believability and engagement, our drama segments are not currently generated automatically, but instead are written by dramatists and recorded by voice actors with drama experience. To reinforce to the group members that they are seeing a drama rather than hearing about dry scientific facts, we import familiar expectations from drama and theatre, such as the bell chimes used to warn of the start of the next performance, animations of red curtains opening and closing on the display screen before and after a performance, and occasionally, applause at the conclusion of the audio drama. Figure 1 shows red curtains opening at the beginning of the presentation; the phone also vibrates at this point to gain attention, but then once the drama has begun, only the audio is active so that the visitor is not distracted.

3. System and Interface

DRAMA^{TRIC} presents audio dramas to small groups of museum visitors using smartphones or tablets as the presentation device (each has a color screen, earphones, WiFi, and various internal sensors such as a compass). Each device communicates with the other devices in the group and with a coordinating server [8], and receives updates once per second for position, orientation, proximity and voice information coming from all the sensor devices belonging to the group. Position is determined by instrumenting the museum with downward-facing WiFi beacons in the ceiling above, while proximity, orientation and voice level are

determined by a small neck-worn device. When this device is within range of a beacon, that beacon's code is included in the device's sensor stream and is forwarded to the server and thus other group presentation devices via existing WiFi routers in the museum.

Knowing visitor position is necessary to initiate the proper drama at the right time, while visitor proximity enables the system to determine the degree of group cohesion. Individual voice activity detection is also important for inferring group conversation, and since the mobile nodes' microphones are pointed upwards towards the wearer's mouth, the difference in intensity of the audio signal between two mobile nodes in close proximity can help determine that conversation is occurring, and its characteristics: who is talking to whom, the length of conversational turns, and who talks most or least. Further inference allows us to assess the effectiveness of any given narrative variation techniques by looking at how much conversation occurred immediately afterwards.

When visitors are walking around the museum and not hearing a presentation, a map of the museum is displayed along with a picture of what they should be seeing around them (Figure 1, left). The visitor's current location on the map is updated whenever the underlying positioning system locks on to them. Once the group arrives at a position with a presentation, the map disappears, the curtains open, and a static drama scene is shown while the audio drama plays.

While group members listen to their audio channel, on their own phone's screen they see a set of images representing the main and supporting characters speaking in that drama, along with information about who or what they and the other group members are hearing. For instance, if visitors Antonio and Emily are listening to a scene on a ship, Emily might be listening to the ship's carpenter while seeing a large picture of the carpenter with a message saying "You are hearing the Carpenter", next to a small picture of the captain with a message under it saying "Antonio is hearing the Captain", while Antonio sees the version appropriate for him. The screen's graphics are deliberately kept simple to enable visitors to quickly understand there are hearing variations, and to allow visitors to spend more time looking at the artifacts or interacting with each other rather than staring at their smartphones.

Visitors hear self-contained drama segments that are combined adaptively and dynamically, and the next variation technique can be algorithmically determined based on sensor data. Our baseline algorithm sums up the voice level readings during a 1-minute "observation period" after drama presentations. If this average is above a threshold we assume that the current technique is working and retain it for the next drama; otherwise it switches technique to see if a new one will increase conversation after the next presentation.



Figure 1. DRAMAtric graphic interface: museum map, start of a drama, and variations for different visitors.

To our knowledge research of group activity has rarely emphasized the mobile aspect. One of the closest works is by Kim and colleagues [7], who used a portable device called the "sociometric badge" to monitor speaking activity and other social signals in a team. Their work does not involve precise positioning and their aim is mainly meant to give feedback to individuals about overall group behavior via a graphical representation of the group behavior on a private display. However, their goal was very different: through reflection on their individual behavior, as represented on the display, participants were shown to better respect speech turns and to reduce overlap when talking. Other localization approaches have been used in specific situations and can be costly [12], or have been based on readily available technology like GPS and are thus more suited for outdoor use [5].

4. Technique Switch Museum Experiment

We wanted to explore the space of potential adaptivity functions by looking at a Technique Switch hypothesis: that changing from one variation technique to another in successive audio drama scenes will have an impact on the amount and/or quality of conversation. Specifically, we chose to test two dramas using the telephone technique, then either continuing with the same technique (Condition 1) or changing to the Blurring technique in the third drama (Condition 2). This allowed us to test the consequences of changing technique with an eye toward exploring what types of adaptivity might be successful.

Changing the narrative technique is per se a very dynamic element of the system, representing a true point of adaptivity. It may for instance help keep visitors' attention high or add to the general level of curiosity, but it may also have unforeseen consequences, such as causing too much confusion. While a variety of

selection algorithms may be used in a final deployment, for this experiment we purposefully started with what corresponds to a naïve hypothesis, namely that if the system determines that a given technique is no longer successful with this group, then it would change to another technique. The Technique Switch hypothesis we chose reflects this simplicity: does switching variation technique at a single point lead to more conversation immediately after that switch?

Before each pair arrived, the system was initialized and the microphone signal was synchronized to the system's sensor data. Each pair was assigned to one of two conditions, either hearing a drama with the telephone technique at all three primary drama positions (Condition 1), or hearing the telephone technique twice followed by the Blurring technique for the third position (Condition 2). Our reason for not switching technique immediately after the first position was to minimize bias stemming from the "surprise" of hearing different audio for the first time (this did in fact occur, with a 6% increase between the first two audio dramas that used the same technique). We did not specifically control for story quality effects, and we allowed pairs to choose their own path and then checked post-hoc for balance.

Subjects were either native or very fluent English speakers, recruited via posters placed around the University of Haifa campus where the Hecht Museum is located, and were paid 50 shekels each (about \$10) for their participation. 29 subjects were between 18 and 30 years old, 8 subjects between 30 and 45, and 3 were older than 45. There were 17 women and 23 men, and 17 of the pairs described their relationship as being friends (1 as classmates, 2 as a couple). The subjects were reasonably balanced across the two conditions for a number of criteria: age, gender, native language, time

Talking by Condition and Drama

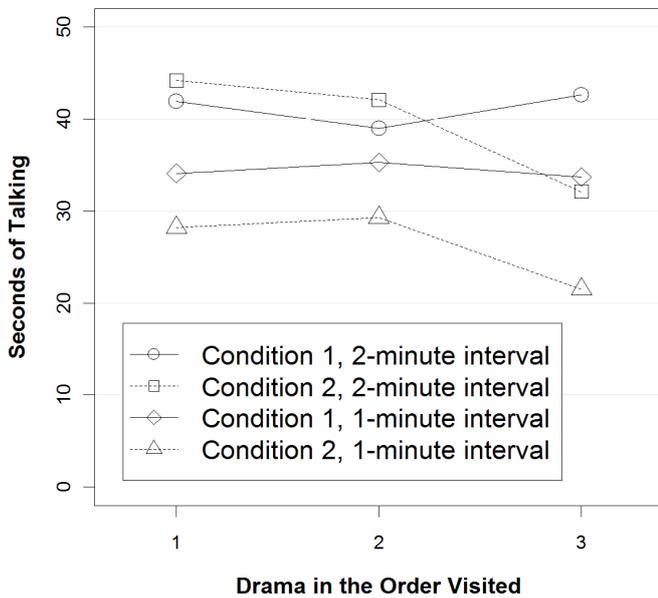


Figure 2. Talking change by condition.

of day of their tour, relationship, the order of exhibits visited, and the last exhibit visited.

After the subjects arrived and gave their consent for participation and data collection, they were shown how to use the device, and how to find their way to the beacon positions using the map on the device. They were explicitly told that they could go through the rooms in any order, that they might not hear everything the characters are saying, and that they might not hear the same thing as each other.

For the experiment 20 pairs of subjects walked through the museum listening to the dramas at each position with a smartphone equipped with earphones. They were given brief training, showing them how to use the device, and how to find their way to the beacon positions using the included map. To determine how much each pair talked, as well as the content of their conversations after the experimental manipulation, we attached a small digital voice recorder to the mobile node worn around their necks.

The resulting experiment data are 20 pairs of audio files, each about 1 hour, and the associated system events (such as initiating playback of an audio drama). Because a prior formative experiment [4] had shown that some subjects talked to each other about the ongoing audio drama while it was still playing (and thus presumably talked less during the observation interval immediately afterwards) we extracted 2 minutes of audio data per drama per pair: from 1 minute before the beginning of the observation

interval through the end of the 1-minute interval. We sent the resulting 2-minute audio files for professional transcription, with later in-house correction where possible. Some subjects who were near-native English speakers spoke in Hebrew, so we translated those portions of their transcripts into English.

The data thus consists of 6,840 total seconds of audio (19 pairs x 3 dramas x 2 minutes), for a total of 951 annotated turns. The average pair talked for 120.8 out of 360 seconds (33.6%, SD=50.7sec) and had 50 turns (SD=24.5). In condition 1, the average pair talked for 122.3 seconds (SD=26.5) with 47 turns (SD=18.0), while in condition 2 the average pair talked for 118.9 seconds (SD=67.2) with 52.5 turns (SD=29.9). Thus conversations in condition 2 showed significantly more variation in length and number of turns than those in condition 1. Frequent overlapping of turns was seen for the most talkative pairs, and only a few of the least talkative pairs had a dominant speaker, as measured by time but not turns. The longest utterance was 25 seconds and the longest dialogue was 36 turns.

To confirm the Technique Switch hypothesis, we needed to look at how much subject pairs talked to each other immediately after hearing the audio drama. We thus measured the amount of talking in 1-second intervals over the 1-minute period in terms of total elapsed time, conversational turns and semantic annotations. The results show that pairs in Condition 2 talk slightly more in the first two dramas, but immediately after in drama 3 there is a striking change: pairs in Condition 1 talk significantly more $\chi^2(1, 1476) = 11.8, p < 0.001$ than in Condition 2 (Figure 2). Thus the general hypothesis was confirmed: switching techniques had a significant impact on the amount of conversation immediately after the switch. Because conversation also occurred while the audio drama was still playing, we also examined the 2-minute period centered at the end of the drama, but there was no change: the amount of talking was still significantly lower after switching technique in the third primary drama segment (32.1 seconds for Condition 2 compared to 42.56 seconds for Condition 1, or a 24.6% decrease, $\chi^2(1, 1476) = 21.2, p < 0.001$). Thus we can confirm that adaptive manipulations can have a significant impact on the amount of group conversation.

We also conducted a semantic conversation analysis of the tours using the following annotation scheme:

- 1: **Silence** – No conversation occurs in the interval.
- 2: **Indecipherable** – Whispering, or face turned away.
- 3: **Irrelevant** – Unrelated to the museum/experiment.
- 4: **Experiment** – Talk about the experiment itself.
- 5: **Technique** – Discussion about narrative variations.
- 6: **Content** – Discussion about the museum exhibits.

As can be seen in Table 1, most conversation is devoted to talking about museum themes, followed by talking about the experiment itself (which decreases

Table 1. Content-categorized talking by condition, with proportion of category given total available time.

| | D1-2 | D1-3 | D1-4 | D1-5 | D1-6 | D2-2 | D2-3 | D2-4 | D2-5 | D2-6 | D3-2 | D3-3 | D3-4 | D3-5 | D3-6 |
|--------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| TOTAL | 23 | 72 | 266 | 156 | 301 | 24 | 136 | 151 | 65 | 396 | 35 | 74 | 179 | 57 | 360 |
| Cond 1 | 6 | 46 | 120 | 55 | 151 | 9 | 63 | 44 | 47 | 188 | 24 | 46 | 73 | 32 | 209 |
| Cond 2 | 17 | 26 | 148 | 101 | 150 | 15 | 73 | 107 | 18 | 208 | 11 | 28 | 106 | 25 | 151 |
| C1 % | 0.3% | 2.0% | 5.2% | 2.4% | 6.6% | 0.4% | 2.7% | 1.9% | 2.0% | 8.2% | 1.0% | 2.0% | 3.2% | 1.4% | 9.1% |
| C2 % | 0.7% | 1.1% | 6.4% | 4.4% | 6.5% | 0.7% | 3.2% | 4.7% | 0.8% | 9.1% | 0.5% | 1.2% | 4.6% | 1.1% | 6.6% |

after the first drama as the pairs become accustomed to the hardware and the procedure). The table shows this data grouped by drama position number (which is variable as each pair chose their own path) where a given drama m and semantic code n is labeled $Dm - n$. Indecipherable talk and talk about unrelated themes remains minimal as the pairs proceed through the museum, providing evidence that pairs were engaged with the audio drama tour. As before, contentful discussion in condition 1 increases, while in condition 2 it rises then falls after the adaptive change in technique.

5. Conclusions and Future Work

We have described the DRAMATRIC system that presents audio dramas for small groups in cultural heritage settings for the purpose of increasing group conversation, which ethnographs have shown leads to more enriching cultural experiences. DRAMATRIC uses museum sensor data such as current visitor position to adaptively modify the audio dramas. Using specific techniques to create narrative variations that different group members hear, we intentionally try to manipulate conversation (and thus social interaction) as a reaction to those variations. We created audio dramas for the Hecht Museum at the University of Haifa, and conducted a full-scale experiment that successfully determined that changing the narrative variation technique could in fact lead to observable differences in group behavior, namely the amount of conversation immediately after an adaptive action. This should lead to specific proposals for promoting social and psychological well-being by changing group dynamics in positive ways.

We are currently extending DRAMATRIC to support remote participation where a family tours the museum while a relative, an elderly shut-in who cannot go to the museum in person, can participate as a group member and remotely experience both the drama tour and what their family members are seeing and saying.

References

- [1] AOKI, P., GRINTER, R., HURST, A., SZYMANSKI, M., THORNTON, J. and WOODRUFF, A. (2002) Sottovoce: Exploring the interplay of conversation and mobile audio spaces. In *Proceedings of ACM SIGCHI*: 431–438.
- [2] BAUM, L., and HUGHES, C. (2001) Ten Years of Evaluating Science Theater at the Museum of Science, Boston. *Curator: The Museum Journal*, 55(4), 355–369.
- [3] CALLAWAY, C., STOCK, O., DEKOVEN, E., NOY, K., CITRON, Y., and DOBRIN, Y. (2011) Mobile drama in an instrumented museum: Inducing group conversation via coordinated narratives. In *Proceedings of the 16th International Conference on Intelligent User Interfaces*. ACM, New York, NY, USA, 73–82.
- [4] CALLAWAY, C., STOCK, O., DEKOVEN, E., NOY, K., CITRON, Y., and DOBRIN, Y. (2012) Mobile Drama in an Instrumented Museum: Inducing Group Conversation via Coordinated Narratives. *New Review of Hypermedia and Multimedia*, 18(1–2), 37–61.
- [5] CHEVERST, K., DAVIES, N., MITCHELL, K., FRIDAY, A., and EFSTRATIOU, C. (2000) Developing a context-aware tourist guide: some issues and experiences. In *Proceedings of the International Conference on Computer-Human Interaction*, ACM New York, NY, USA, The Hague, Netherlands, 17–24.
- [6] DAMIANO, R., GENA, C., LOMBARDO, V., NUNNARI, F. and PIZZO, A. (2008) A stroll with Carletto: adaptation in drama-based tours with virtual characters. *UMUAI*.
- [7] KIM, T.J., CHANG, A., HOLLAND, L., and PENTLAND, A.S. (2008) Meeting Mediator: Enhancing Group Collaboration with Sociometric Feedback, In *Proceedings of the ACM Conference on Computer-Human Interaction*, Florence, 3183–3188.
- [8] KUFLIK, T., STOCK, O., ZANCANARO, M., GORFINKEL, A., JBARA, S., KATS, S., SHEIDIN, J. *et al.* (2011) A visitor’s guide in an “active museum”: Presentations, communications, and reflection. *Computing and Cultural Heritage of the ACM* 3(3).
- [9] LEINHARDT, G. and KNUTSON, K. (2004) *Listening in on Museum Conversations* (Altamira Press).
- [10] PETRELLI, D. and NOT, E. (2005) User-centred design of flexible hypermedia for a mobile guide: Reflections on the hyperaudio experience. *UMUAI*.
- [11] STOCK, O., ZANCANARO, M., Busetta, P., CALLAWAY, C., KRÜGER, A., KRUPPA, M., KUFLIK, T. *et al.* (2007) Adaptive, intelligent presentation of information for the museum visitor in PEACH. *UMUAI*.
- [12] ZIMMERMANN, A., and LORENZ, A. (2008) LISTEN: a user-adaptive audio-augmented museum guide. *User Modeling and User-Adapted Interaction*, 18(5), 389–416.