Medium as frame: comparing mobile audio and video interactions in informal learning contexts

Brett Oppegaard*

School of Communications,
University of Hawaii,
550 Campus Road, Crawford 330,
Honolulu, 96822-2217 HI, Hawaii
Email: brett.oppegaard@hawaii.edu
*Corresponding author

Michael Rabby

The Creative Media and Digital Culture Program,
Washington State University Vancouver,
14204 N.E. Salmon Creek Drive,
Vancouver, 98686 WA, USA
Email: mkrabby1@gmail.com

Abstract: Concept-driven interaction design opens new pathways for research of information and communication technologies (ICTs) in educational contexts, between the particulars of a case study and the abstractions of theory. Explorative research in this project tested foundational theoretical ideas, such as medium specificity, through concrete designs in an authentic setting. During these separate but similar procedures, conducted a year apart, ordinary users were given representative tasks on mobile devices in order to examine the levels of involvement, social facilitation, and satisfaction generated by differing media forms within the mobile delivery system. From this perspective, direct comparisons could be made, for example, between audio and video forms. Enabling such comparisons has grown in importance with the advent of mobile and other convergent technologies that blend mediums to bring together media organisms to comingle. In the case of a smartphone, for example, mobile media designers now can choose which medium (text, audio, video, animation, etc.) they want to use within their master medium, which adds to the complexity of the design endeavour but also to the potential for new integrated and interactive forms to emerge as well as for more mindful context-tailored solutions.

Keywords: medium specificity; medium studies; mobile media; information and communication technologies; ICTs; smartphone; audio-video comparison; involvement; social facilitation; satisfaction; medium layering.

Biographical notes: Brett Oppegaard is an Assistant Professor at University of Hawaii. He studies ubiquitous computing and mobile media. He was the individual recipient of the regional and national 2012 George and Helen Hartzog Award for his research into mobile app development and media

delivery systems within the National Park Service as well as the national 2013 John Wesley Powell Prize winner for outstanding achievement in the field of historical displays. He was chosen for a National Endowment for the Arts fellowship as a journalist and also has earned National Endowment for the Humanities grants as a scholar for his innovative mobile media research projects. Those projects include collaborations with the USA's first national park, Yellowstone, and the National Park Service's Harpers Ferry Center, the Interpretive Design Center of the federal agency.

Michael Rabby began researching how people talked to each other over email in 1995, and has continued investigating various facets of internet relationships ever since. He was one of the first researchers to study the actual content of email messages, the subject of his Master's thesis, and to use the internet as a tool to collect survey data for his PhD. This research path has led him through areas such as mobile apps, online and offline romantic relationships, online relationship maintenance, online impression management and self-presentation, the use of social media by deployed troops, and adoption of tablets for journalism.

1 Introduction

Researchers for decades have speculated and experimented with mobile devices as both formal and informal learning technologies, but in the past few years, those inquiries widely have spread into authentic educational contexts (Crompton, 2014; Kay and Goldberg, 1997; Taylor et al., 2006). As part of such investigations, some scholars have been exploring ways in which the medium matters, by comparing mobile technologies to analog or digital counterparts. They have looked at how the medium within the medium matters, making intradevice comparisons (Fagerjord, 2011; Kang and Gretzel, 2012; Maniar et al., 2008; Novey and Hall, 2007; Piper et al., 2012; Jasim and AbuSeileek, 2015; Soo et al., 2014; Suh et al., 2011). In the latter cases, researchers are confronted with unprecedented compositional complexities to consider. These at first might seem intricately intermingled, inaccessible, and irreducible. The emerging mobile medium – differentiated by bringing together media of all types, delivered through location-aware smartphones, tablets, watches, etc. – essentially absorbs its predecessors in an ultimate act of convergence. Yet the mobile medium also interacts in novel ways with the physical world – including with people, objects, and settings – in multiple and overlapping real-time dynamics. This complex amalgamation of place, mobility, information, and technology, fuelled by sociability motivations, has been generating the first distinctive new forms of 21st century media (Hjorth, 2013). As such, the mobile medium clearly has prodigious learning-technology potential, but it also remains shrouded in many deep mysteries. During its emergence of the past decade, the mobile medium has been noted by scholars throughout this paper for the ways in which it has been changing individuals and societies, presenting countless intellectual puzzles for scholars. What has been less obvious and less recognised – and only lightly studied – in this process is that the mobile medium also has been presenting opportunities to apply powerful new tools for addressing these related learning issues, if we only would employ the mobile medium itself to study mobile issues.

The media-ecology matters of the convergent media forms (such as texts, audio, animation, still-image photography, and video) swirling within the mobile medium have been mostly sidestepped to this point due to a variety of practical and theoretical reasons. The complexity of this topic can make it difficult to discuss and isolate issues in meaningful ways. However, Stolterman and Wiberg (2010) provided a powerful point of orientation for this type of inquiry through their argument for a middle ground of mobile media research. This middle ground, or 'intermediate level', as we will call it – in the zone between detailed case studies and abstract theorisations – allows explorative research to manifest theoretical ideas in concrete designs, valuing prototyping as a form of scholarly inquiry. They label this approach *concept-driven interaction design*, as it is both conceptually and historically grounded but also guided by theoretical considerations.

Following that methodological lead into the intermediate medium-within-the-medium layer, we sought a better understanding of the intrinsic differences between mobile audio and mobile video in informal learning contexts. This research study has been carefully designed to reveal those nuanced ways in which people experienced those media forms, through mobile technologies. To test assumptions made about those experiences, from such a perspective, these types of inquiries can challenge learning-technology research in intentional ways. They can do this without getting too finely detailed or too broadly nebulous by departing from a conceptual base, by exploring through hands-on design and development of artifacts, and by investigating a conceptual problem rather than a practical and situated one. This generative 'intermediate-level' knowledge is more abstracted than particular instances but does not aspire to the generality of theory (Höök and Löwgren, 2012). Instead, the particulars provide grounded anchors, and, in conjunction, the theory establishes the parameters of the knowledge space. This generative sensemaking process generally operates in *hoverspace*, above the grounded particulars and out of the clouds of abstraction.

This study therefore attempted to advance conceptual understandings about the mobile medium beyond use-case scenarios, without ambitions of significant alterations to medium specificity theories. Through this approach, we take a conceptual look at the in-between realms, where situational context is not highly critical, and theoretical abstraction is not highly valuable, where the medium choice is not predetermined by extremes of context or setting or personalities but empirical data. We also describe how we built mobile apps as research instruments to be tested in authentic learning situations with real audiences in real settings, transparently displaying the interworking of the experiments. Specifically, we sought to determine if audio media would be proven through empirical evidence to have a cognitive-resonance edge over video media when used for place-based informal learning, per general assumptions in the site interpretation industry, where audio tours are common and have been prominent for decades.

In that pursuit, we start this paper with a literature review of scholarship about place-based interpretive media in learning environments, followed by the conceptual development of the medium as a frame, and how that concept can be applied, in this case, through a form of inquiry that included mobile app prototyping in authentic contexts. Our method of inquiry is described, in detail, including the prototype designs, the sampling process, and the research procedures. Our findings then are discussed, acknowledging limitations and projecting potential paths for future exploration.

2   From wall text to tablet computers

At the core of the conversation about this complicated topic is an attempt to make direct comparisons between audio and video forms at a place-based attraction. While many factors affect decisions about content development and delivery form, the intent here is to at least move the discussion past the idea that adding mobile content – in any form – to a mediated learning environment enriches it. That proposition seems evident enough as to not to expend energy exploring it. Beyond simply enhancing the visitor experience in some amorphous way, as Falk and Dierking (2008) argue, new technologies need to improve the visitor experience better than competing technologies and in ways that are cost-effective. That is the survival-of-the-fittest component inherent in our inquiry. An investigation therefore is warranted that moves beyond historical and commercial rationales to attempt to determine if better options exist for informal learning at place-based attractions, or if audio forms truly are inherently superior to other forms, such as video, in ever-evolving ubiquitous computing environments.

Using historical examples as representative illustrations, one can trace the evolution of place-oriented media forms through their various delivery systems into the realm of informal learning technologies. In the early stages, for example, museum curators simply provided their audiences with wall text. That context-changing informal learning innovation allowed organisational spokespeople to leave their assigned posts, near important places, but also to continue to shape the course of discussion around their prized attractions. Later technological advances included expanded object labels (sharing detailed artist and contextual information), followed by the development of handout analog literature, such as leaflets and brochures, and eventually the creation of the audio tour. As a mediated extension of the in-place lecturer, audio tours pervasively could channel the museum's voice into the ears of its visitors as they stood in front of an object, thereby liberating one's eyes to actually see it (Samis, 2008; Tallon, 2008).

Two companies stabilised and patented the audio tour market, starting with Acoustiguide, in the late 1950s, and followed by Antenna International in the 1980s. As Tallon (2008) noted, these companies robustly built their businesses, as the Coke-Pepsi of audio tours, with a virtual lock on the market, and with a captive clientele. Yet, as a consequence, public research in this area dissipated as part of their combined push for subcontracts of proprietary control systems. Studies about the visitor experience with audio tours, in turn, became mostly cursory, and, as Tallon (2008) found, those studies typically focused upon purchasing statistics. The medium, from this viewpoint, became more about the profit-generating possibilities than the ideal mode of interpretation. The 'strong concept' of social navigation, as summarised by Höök and Löwgren (2012), can be viewed as another example of intermediate-level knowledge, and another way to look at this particular situation, in which some learning technology decisions appear based on the decisions of others, and neither the particulars of a case, nor a grand theory, thoroughly can explain the situation.

One of the founding scholars in mobile place-based media, Jeremy Hight, for example, recalled how he and his research team considered creating layers of visual media to accompany audio media in their seminal mobile narratives in the early to mid-2000s but found it important to stick to audio alone, because it allowed a sense of experiential blending and overlapping (Miranda, 2008). As Hight (2005, 2006) had noted

previously, the effect is of being in one place and in one time, while the augmentation allows the past and unseen to come to life. Visuals compete, and if the information is primarily text on a device, it becomes separated from the physical world and secondary to it. Crucially, this can have the effect of making place and the data of place two things independent of each other and not integrated into juxtaposition and comparison in the same way. As another example of this acoustic-oriented approach, Dow et al. (2005) chose to deliver audio-only media, rather than video, in their research prototype, as they felt visual media would distract the user's attention from their surroundings. Such thoughts about medium choice can be concluded from many different perspectives, from a response to the particulars of a situated case to the abstraction of a grand theory, with debate about the significance of each, and the appropriate metrics. An inherent tension, for example, exists from the resource-allocation perspective, of what is valuable, in terms of interpretation media, and, what costs justify the expenses of producing, developing, and maintaining those artifacts or channels. Kang and Gretzel (2012) note the theoretical proposition that human voices convey rich social information, and, therefore, audio tours could add richness and value to tourist experiences.

Yet this effectiveness depends on what is measured, and how, with mixed results. Novey and Hall (2007) cited studies that found although audio components attract attention, they did not necessarily sustain that attention longer than nonauditory exhibits (e.g., Beer, 1987), nor did they have a substantial effect on information exchange (e.g., Peart, 1984). Novey and Hall's own research demonstrated insignificant learning gains, compared to what visitors achieved through using signs and other analog media. Suh et al. (2011) noted evidence that audio-only systems failed to keep users visually stimulated as well. Kang and Gretzel (2012) suggested that mobile technologies increasingly mediate tourism experiences, and those audio forms, such as podcasting, have advantages and great potential to tourists and tourism organisations, based on the rich social information conveyed by human voices. They did not note the distinction that other mobile media forms, such as video and animation, also could include human voice. Fagerjord (2011) further argued that sound is more effective than visual modes as it allows the audience to look around while listening. Maniar et al. (2008), in contrast, described and defined contexts in which video-based media in mobile (m)-learning situations could be more effective than learning from text or audio, by comparing the forms. Yet overall, each of these studies has had limits in scope; consequently, more empirical evidence about medium comparisons within mobile contexts is needed, hence our attempt to approach this issue from a new angle. In summary, audio tours were developed, and continue to be developed – per the strong concept of social navigation – not because they necessarily have been proven to be more effective at visitor engagement than other media forms. This dynamic has occurred regardless of any novel openings for innovation afforded by mobile technologies, such as the ability to bring high-quality video screens into places they have never been able show visual media before, such as in the middle of a Gettysburg battlefield, or at the base of Old Faithful in Yellowstone National Park.

3   The medium as a frame

In a content-rich place, like a US National Park, people explore their environment by attending to a miniscule portion of all of the information, or data, or interpreted signals of

events, available at any given moment (Klein et al., 2007). This data can come in many forms, including sensory and mediated, and can be parsed in a variety of ways. In the process of making sense of the perceived and relevant data, i.e., learning, people use what can be considered a perspective, or a viewpoint, or a framework to define and shape that data; in turn, these so-called frames also can take many forms, including stories, scripts, plans, maps, diagrams, audio clips, and videos (Klein et al., 2006a, 2006b; Klein et al., 2007; Pirolli and Russell, 2011). Frames help both the designer and the audience to manage attention while also defining, connecting, and filtering the data according to the frame. Frame, framing, and framework are common terms outside of formal scholarly discourse; Entman (1993), in response, criticised the fractured and casual nature of the terms as often used loosely in academia, suggesting the focus of a frame remained on two crucial elements: selection and salience. With those in mind, a designer selects some aspects of a perceived reality to make them more salient in the communicating media form, to promote a particular problem definition, causal interpretation, moral evaluation, and/or treatment recommendation. In the context of this study, the term frame is used as a frame of communication, which includes words, images, phrases and presentation styles, such as data framed and delivered via audio or video (Druckman, 2001; Gitlin, 1980).

When a new medium emerges, such as the mobile medium, historical and cultural influences, in hybrid understandings, come along as well, as older media gets blended into new forms, in a process Bolter and Grusin (1998) describe as remediation. Richardson (2011) connected the concepts of remediation and convergence (Jenkins, 2006) through medium specificity theory by noting that each interface, even when experiencing different kinds of services and content within a single apparatus, such as a mobile device, can be interpreted in terms of specific and differential effects, each demanding a particular mode of embodied interaction. From this perspective, an audio frame and a video frame arguably can be compared as long as the data inside the frame is selected to be salient and considered equivalent, in terms of media richness. This key conceptual viewpoint provided the intermediate-level platform upon which interaction with a mobile prototype could be examined to study general design implications and to challenge an assumption that acoustic media is inherently better than video media for informal learning at place-based attractions (Dow et al., 2005; Hight, personal communication, 2013). Identifying exactly what has, or is, changing in the media ecosystem with the ubiquity of mobile technologies today can be difficult to pinpoint. Through this comparison of frames, this study does not attempt to compare wall panels or wayside signs to digital forms. Instead, it specifically examines the differences, or lack of differences, in the responses users have to carefully crafted media forms (audio/video/audio-video), through the same type of digital device in the same physical place at roughly the same time, with the same external experiences, as a way to isolate the frame comparison but also to operate within a realistic setting in a real informal learning situation.

A further goal of this work is to use the frame comparisons to create empirical measures of the frames, and to see how they perform, against established metrics and against each other, without the historical baggage tainting such comparisons. By examining these frames carefully from a conceptual perspective, the mediated interactions come into focus, along with the effects that select media forms have on visitors to a place-based attraction. Instead of a content analysis or measurement of media richness, this study examines the informal learning aspects involved with the interpretation of place, from an intermediate-level perspective focused on the medium

within the medium (the audio or video within the mobile medium), to determine conceptually how media changes its potential when its access and delivery methods change.

4   Exploring engagement through prototyping

Using this explorative concept-driven approach, guided by Stolterman and Wiberg's (2010) descriptions of related methodological activities, we created, tested, and refined mobile app prototypes multiple times in an effort to create field research instruments suitable for such an inquiry into this intermediate-level medium debate. Particular attention was paid to the nuances of the data used to carefully construct both the audio and video frames. Key concerns included equal length, equivalent information richness, and general comparability of the clips. We edited the video piece first in this process, with the comparability concerns at the forefront of the mind, and then pulled the audio track directly from the video, to provide the most comparable form of media possible. The video piece did have some visual data inaccessible to the audio-only audience, such as the look of the musicians and the archaeologist in action, but we felt those parts of the video also were of nominal experiential value that would be compensated for in audio through the increased cognitive focus on the sounds of the voice, music, and instruments. Because this perceived equivalence by us is at the core of the argument for comparability in this experiment, hyperlinks are provided to allow for independent analysis of these claims and efforts. The final video clip can be viewed here: https://goo.gl/t3LnVq. The soundtrack is exactly the same for the audio-only clip, a transcript of which can be read here: http://goo.gl/TACBqK. It also can be heard through the provided video link. The point needs to be made here that this was not a simple stripping away of audio from a video; it involved a crafting of an audiovisual piece that had equivalent media richness whether the audience member listened to it or watched it.

During this meticulous research-instrument crafting process, with a focus on the 'intermediate-level' of knowledge, we determined that it would be better to compare a roughly similar audio and video clip, versus the polarising alternative (to make the best possible video for the situation and the best possible audio and then compare them at the strongest points in this one particular situation). In our approach, we decided to put our emphasis on a regulated representation of each medium, in which the data within the frame was as close to the same as possible, allowing a closer examination of the frames than the data. This design provided a clear means to explore the engagement levels of the user from a variety of different perspectives, including involvement and social facilitation.

One means of examining the effectiveness of a particular medium involves exploring the level of engagement users have with it. Mersey et al. (2010) define engagement as the collective experiences that readers or viewers have with a media brand, when experiences are considered a specific set of beliefs that consumers have about how some media brand fits into their lives. To consider the intent and understand the effects of such experiences, they contend, is an essential first step in developing and positioning a successful media brand.

O'Brien and Toms (2008) noted the complexities of the concept of engagement in their attempt to define a conceptual framework for people's' experiences with technology. Their research indicated ten attributes of engagement for people when interacting with an

object: aesthetic and sensory appeal, attention, awareness, control, interactivity, novelty, challenge, feedback, interest, and positive affect. With mobile apps, this increased level of engagement is apparent. The user's ability to interact with a media object moves beyond listening, expanded to include actions beyond stopping, rewinding and fast-forwarding. Now, the user can experience different types of media, choose a variety of options and activities, explore the environments around them using augmented reality, record their own experiences at the site, and share these experiences.

Engagement therefore represents an important measure of a user's connection with a mobile app in this study. In the past, measuring engagement could best be explored using a post-hoc test to measure learning, with perhaps further testing to include time spent in an exhibit. Here, we use two aspects of engagement as a means to observe the influence that different media might have for users of a mobile app. First, feeling of involvement describes the feeling of being drawn into and involved in a task, and their overall assessment of the experience as 'fun' (O'Brien and Toms, 2010). The second type of engagement, social facilitation, explores the extent to which participants feel compelled to discuss what they experienced with other people. It indicates if people were captivated enough by something that they had learned to share it with their social circle, in a sense measuring engagement beyond the initial connection, through the processing and fulfilment stages.

Investigating the effects of these different types of media (audio-only, video-only, and audio/video) in a mobile app based on users' engagement and happiness with a media object leads us to the following research questions:

RQ1 Do different types of media within a mobile app lead to a greater feeling of involvement with the app?

RQ2 Do different types of media within a mobile app lead to greater urges for social facilitation?

5   Method

5.1   Study 1 procedures

The first round of data collection occurred on October 6, 2012, during the Old Apple Tree Festival. The free annual event brings together the people of Vancouver, WA, and the surrounding communities for an afternoon of music, food, and demonstrations, with the ultimate purpose of honouring the Old Apple Tree. The tree, which dates back to 1826, represents the last aboveground remnant of the historical place, near the Fort Vancouver National Historic Site, of what once was a large grove of fruit trees. The tree now is the centrepiece of a small and isolated urban park, surrounded by highway and railroad infrastructure. The park is seldom-used most of the year but attracts several hundred people to this particular event. At the start of the festivities, we planted two signs at the front of the tree that promoted and informed people about the research project. The two co-authors stationed themselves next to the Old Apple Tree and solicited people as they passed by and/or stopped at the tree, trying to gather as many participants as possible. One of the researchers used a smartphone, and the other used a tablet computer with a connection tethered to the smartphone. No significant differences emerged between the tablet and smartphone conditions on any of the tested variables.

Participants were asked if they would like to use a new mobile app to learn about the Old Apple Tree. They were told the app would take between 3–5 minutes (average time on the app = 6.81 minutes, with a median of 7 minutes). Participants were randomised into one of three conditions that differed based on the media viewed: audio-only, video-only (captioned with a transcript of the audio component), and an integrated audio-video presentation. We used the aforementioned media pieces produced specifically for this app. The user was given the media from just one of those conditions and not allowed to participate twice. After the media played, the app invited the user to take a picture next to the Old Apple Tree. If the person opted to take the picture, the person was then given the option to have the picture e-mailed to an account of choice, shared via a Flickr account (a social network that focuses on photos), or not shared at all. These secondary tasks were designed as additional measures of social facilitation. Finally, the app prompted people to complete a brief, 9-item survey, if they so desired; they also had the option to skip the survey completely. The app randomised the questions, eliminating any systematic order effects.

5.1.1 Sample

Collecting data in this manner allowed for a high-quality sample representative of the community. A total of 54 people participated in the app project. Of those, demographic data is available on 41 people (75%); the other 13 people (25%) elected not to complete the demographic portion of the survey. The ages of the participants were split between 15–18 years-old (15%), 19–30 years-old (20%), 31–40 years-old (9%), 41–50 years-old (30%), and 51 years and older (2%). Males comprised 28% of the sample ($n = 15$), and females comprised 46% of the sample ($n = 25$). One person did not specify a gender. Finally, 48% ($n = 26$) described themselves as European American. The categories of African American, Hispanic/Mexican, and Pacific Rim each had one person, with 11% ($n = 6$) describing themselves as 'other', and 6% ($n = 3$) preferring not to identify themselves by race.

5.1.2 Measures

To measure involvement, we used both survey data and behaviour within the app. The survey items came from a subset of O'Brien and Toms' (2010) scale of user engagement. Here, we used three items comprising the involvement variable, measured on a 1–5 Likert scale: 'I felt interested in this app', 'I was really drawn into this app', and 'this app was fun' ($\alpha = .79$). In addition to the survey data, we also took note of people's behaviours while using the app. We operationalised involvement by noting if they chose to take a picture with the tree, as well as if they completed the survey. Likewise, social facilitation was measured in two ways. First, three slightly-modified items from Mersey et al. (2010) online engagement survey comprised the social facilitation variable, measured on a 1–5 Likert scale: 'I will bring up things I saw on this app in conversation with other people', 'This app gave me something to think about', and 'I will use things from this app in discussions or arguments with people I know' ($\alpha = .91$). Sharing the photo (via Flickr or email) was considered means of social facilitation as well, and the behaviour was treated accordingly.

5.1.3 Results

RQ1 investigated the connection between media type and involvement. Table 1 shows the means for each media type. An ANOVA revealed no differences between the media types ($F = .19$, $df = 2, 37$, $p < .83$).

For the measures of behaviour, a chi-square test of the differences between media type and taking the photo revealed no significant differences, χ^2 (2, $N = 54$) $= .44$, $p < .80$, nor did a test of the differences between media type and taking the survey, χ^2 (2, $N = 54$) $= .05$, $p < .98$.

RQ2 investigated the connection between media type and social facilitation. Table 1 shows the means for each media type. An ANOVA revealed no differences in the relationship between media type and social facilitation ($F = .41$, $df = 2, 37$, $p < .67$).

Table 1   Involvement and social facilitation by media type

	Involvement			Social facilitation		
	N	*Mean*	*SD*	*N*	*Mean*	*SD*
Audio	14	4.10	.82	14	3.81	.93
Video	12	4.25	.73	12	4.08	.91
Audio/video	14	4.23	.56	14	3.81	.79

We also used the behaviour of sharing a photograph to measure social facilitation. Of the 32 people who took a photograph with the tree, 24 shared their photos via email, four shared their photos on the Flickr site, and four did not share their picture at all. For the observed measurements of behaviour, a Chi-Square test revealed no differences between media type and sharing the photo, χ^2 (4, $N = 54$) $= .85$, $p < .93$.

Using survey items as well as observation of behaviour through the app, the presumed superiority of audio did not materialise, at least based on the variables tested. In short, media type did not have any significant effect on engagement, a finding that could lead to a variety of interpretations, which will be discussed later in this piece.

5.2 Study 2

As the desired consumption outcome, satisfaction binds many of these engagement factors together into what can be a binary reaction by the audience; a person is either satisfied or not with the media object. Satisfaction is considered central to media use, as organisations and service providers with satisfied audiences likely remain loyal while dissatisfied audiences defect to alternatives. In addition, satisfied audiences likely come back to the same source for more media offerings and also to share their positive experiences with others, as another form of social facilitation (Jacobs, 1999; Patwardhan et al., 2011). Therefore, an examination of satisfaction has become a critical outcome variable for various media objects. Media satisfaction, in turn, comprises a positive general feeling, of varying intensity, evoked by a user's favourable post-consumption evaluation of a medium, media genre, media program, media content, or media-generated activity (Patwardhan et al., 2011). To further validate findings from the previous study, a second study expanded and improved the first study in a few key ways during the next year's festival, in 2013. The app was altered to address navigation and user-experience concerns from the first study, the collection was improved, and a key variable,

satisfaction, was added to the survey. This added a third research question alongside the previously noted questions:

RQ3 Do different types of media within a mobile app improve reported satisfaction with a mobile app?

5.2.1 Procedures

The second round of data collection occurred on October 5, 2013, again during the annual Old Apple Tree Festival (Figure 1). Procedures remained the same as the first study, with a few key changes. First, rather than relying on a mobile network connection to share audio and video data with users, app contents were uploaded into the device beforehand, to lessen technical issues with delivery of the data. When devices failed, they typically worked again after a reboot, and we no longer had to deal with network connectivity issues. This issue rarely surfaced when testing mobile apps in the field and made the data collection reliable and steady. At the same time, it simulated the ability to connect people to the subject and the location. Second, rather than one smartphone and one tablet being used to collect data, six identical Google Nexus 10 tablets were employed, allowing multiple data collection points at one time, crucial at an event that featured ebbs and flows of people. In response to feedback and observations that the media in the first iteration of the app was too long, it was shortened from three minutes to one. This shortened the overall testing time by about two minutes. Rather than connecting to Flickr, the people who elected to share their photos publically were published on the Old Apple Tree's Facebook page. Finally, two additional research assistants helped to collect the data, enabling a greater number of potential participants who could explore the app. The number of participants tripled (from 54 to 169), so this made a worthwhile difference. Guided by the results from the previous study, the app only featured two types of media, audio only and audio/video. The video-only condition was deemed somewhat rare in real-world application and did not demonstrate a meaningful difference from the other media types in the first study, so it was eliminated.

Figure 1 The Old Apple Tree Festival data-collection scene in 2013 (see online version for colours)

Notes: The Old Apple Tree, on the right, served as the centrepiece of the collection process. A principal investigator is shown near the rock in front of the tree, on the left, with a tablet computer, in a tan vest.

5.2.2 Sample

A total of 169 people participated in the 2013 research. Of those, demographic data is available on 115 people (68%). The ages of the participants were split between 15–18 years-old (4%), 19–30 years-old (15%), 31–40 years-old (17%), 41–50 years-old (12%), and 51 years and older (18%); 2% chose not to answer. Males comprised 23% of the sample ($n = 38$), and females comprised 43% of the sample ($n = 73$). Four people did not specify a gender. Finally, 46% ($n = 78$) described themselves as European American, 4% ($n = 7$) as Hispanic/Mexican, 2% ($n = 3$) as Pacific Rim, and one person identified as Middle Eastern. 9% ($n = 15$) described themselves as other, and 7% ($n = 11$) preferred not to respond.

5.2.3 Operationalisation

The survey featured the same measures of involvement and social facilitation. In addition, satisfaction was measured using a modification of Patwardhan et al. (2011) four-item media satisfaction scale: 'using the app gave me pleasure', 'this app was enjoyable', 'I felt good after using the app', and 'I felt happy after spending time with this app' ($\alpha = .88$).

5.2.4 Results

Overall, results in the second study echoed the results from the first study. No significant differences existed between the conditions, even with eliminating one of the media conditions, adding a new variable, and tripling the sample size. The two measures of engagement, involvement and social facilitation, had no significant differences between audio and audio/video. For involvement, a *t* test indicated no differences between the media types, $t(113) =.526$, $p < .60$. Table 2 shows the means and standard deviations for each type. Social facilitation featured a similar result, $t(113) =.168$, $p < .87$.

Table 2   Involvement, social facilitation, and satisfaction by media type

	Involvement			*Social facilitation*			*Satisfaction*		
	N	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*
Audio	55	3.96	.91	55	3.87	.91	55	4.23	.89
Audio/video	60	4.05	.71	60	3.91	.80	60	4.14	.70

Study 2 added the variable of satisfaction, but it did not show a significant difference between the two media types, $t(113) =.606$, $p < .55$. It was the only variable to have higher means for the audio group, but the difference was negligible.

For the measures of behaviour within the app, a chi-square test of the differences between media type and taking the photo revealed no significant differences, $\chi^2 (2, N = 169) = .895$, $p < .34$. Of the 169 people who completed the app, 75 took a picture. Table 3 shows the specific breakdown of where the photos were shared – via email (private) or Facebook (public). It did not have significant differences, $\chi^2 (2, N = 169) = 2.81$, $p < .245$. Finally, media type did not influence whether or not people completed the survey, $\chi^2 (2, N = 169) = .002$, $p < .969$. In total, 115 people completed the survey, versus 54 who did not.
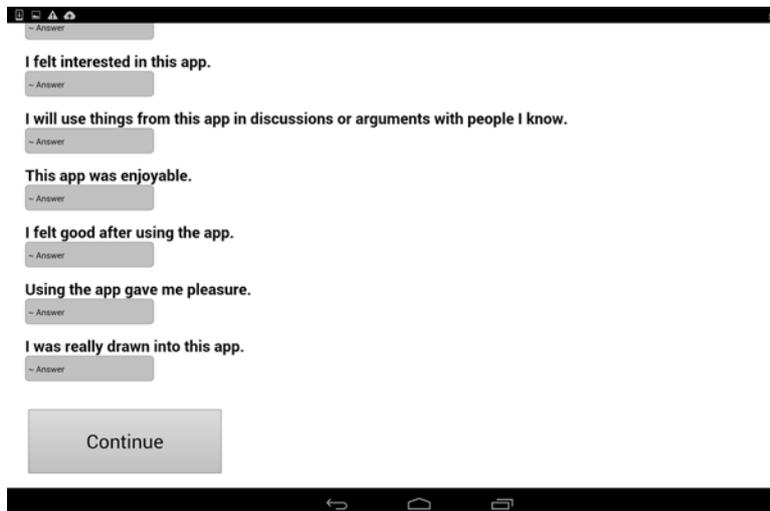
Table 3   Photo shared or not

	Shared via email	*Shared via Facebook*	*Not shared*
Audio	23	14	44
Audio/video	24	8	56

Figure 2   Screenshot of the Old Apple Tree app (a) video prompt (b) survey questions (c) survey choices (d) photo-taking option (see online version for colours)
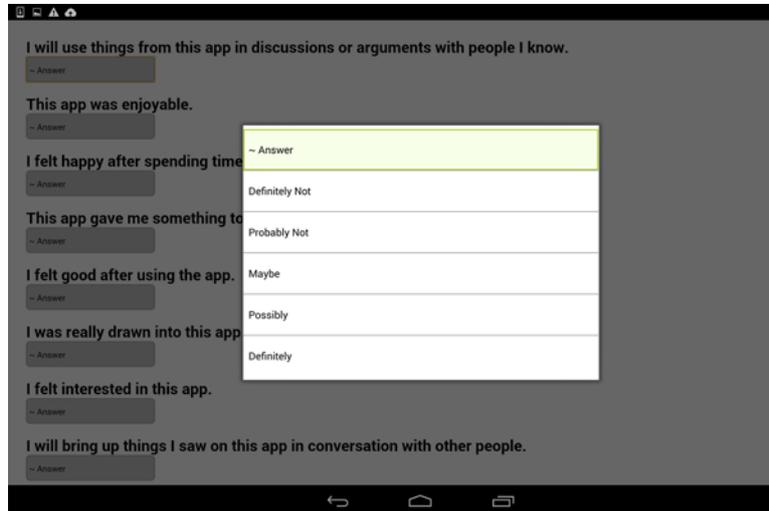
(a)

(b)

Figure 2   Screenshot of the Old Apple Tree app (a) video prompt (b) survey questions (c) survey choices (d) photo-taking option (see online version for colours) (continued)

(c)

(d)

6 Discussion

Mediated interpretation is the core of informal learning and the primary communication channel between organisations and visitors at most place-based attractions today. That relationship between place-based learners and place-tailored media likely will become more significant in the future, as support technologies improve, as media literacy increases, as the 'internet of things' concept expands, and as expectations for better mediated interactions rise accordingly. To improve the design of interactive and informal

learning media in mobile-dominant ubiquitous computing environments, though, much more needs to be known about the ways in which converged mediums operate and engage users within the mobile medium. Between analog (e.g., wall text, brochures) and digital information competing through a thickening information fog, designers need to find a way to better serve their audiences. Through two separate studies using similar procedures, strong and diverse samples, that took into account both survey measures and behaviours performed within the app, we found a consistent result. This has several potential implications for design research into medium specificity issues.

This concept-driven study, for example, suggests that audio tours no longer should hold dominion over mediated informal learning situations, and be thought of as the ultimate extension of interactive and interpretive media. Mobile devices right now can deliver an array of media forms – including video, animation, text, still imagery, and audio – in complex multimedia packages that raise innumerable questions about what works best, when, and why. This study peeled off just one small portion of that intertwined ball of curiosities, in an attempt to determine if audio truly is a superior mobile medium, as some scholars and practitioners seem to suggest (Dow et al., 2005; Hight, 2005; Miranda, 2008).

From the intermediate-level perspective of this study, audio performed no better than video, or even video with captions and no sound, indicating that the playing field for media in the medium might be levelled, by the mobile channelling, and fertile for new exploration. These findings are not intended to imply audio and video will generate equal engagement in particular situations that play to their respective strengths, or against their weaknesses, like in a noisy exhibition hall, or in a quiet museum gallery already filled with visual imagery. These findings do not suggest that just because video might perform as well as audio, pragmatic production concerns should be forgotten, or that the audio-video ratio of informal learning, interactive, and interpretive media suddenly will (or should) surge to a 50–50 ratio. Practices often change slowly, and they are affected by culture as well as possibilities.

What this study should suggest is that audio, as an interactive prompt, and as an interpretive medium in informal-learning situations, does not hold any inherent advantage over video from an intermediate-level knowledge position. With a roughly comparable artifact, mindfully composed of data that does not privilege one medium over the other, an audio frame can be compared to a video frame to better understand ubiquitous-computing environments enriched by mobile media.

As a tangential finding, this study demonstrated that the public could be hesitant to participate through mobile technologies in public forums in which their images would become a part of a communal tapestry. Sharing within a well-understood and controlled circuit, such as via email, happened more often than via a public site, which could prompt a worthwhile follow-up study on related engagement and participation issues.

From the concept-driven interaction design perspective of this study, the media form within the mobile device does not appear to affect engagement or social facilitation, raising intriguing questions about what exactly is the medium in this case, or if the medium, in the mobile delivery system, does not matter as much as it does in other systems. We carefully crafted the media forms in this study to be equal in their information value. Finding ways to compare media forms should be a primary pursuit of media and technology scholars, particularly in a converged environment, when designers easily can make choices between the media forms within the same delivery device, such as a smartphone or tablet computer. How did this effort at such a comparison fare? There

is little else of this sort to compare it to, so these results are intended to start a discussion rather than end it. Despite concerns, including our own, about how pure of a comparison between media forms can be made, even through a conceptual approach using frames, our perspective is that getting as close as possible is better than not trying at all. One advantage of this concept-driven frame approach is that it eliminates many of the primary factors that prevent comparisons from being made, such as production costs, artistic preferences, situational differences, etc. We think this is an approach that could be built upon to create deeper understandings about media forms that before convergent devices were essentially from different worlds, and practically incomparable.

To elaborate, we think that blunt media form comparisons would be weak if they failed to acknowledge and account for the fundamental differences that could underlie the forms, whether that be related to budgets, technical expertise, or aesthetic vision. There would be little point, for example, in comparing the effects of a carefully constructed $100 million Hollywood blockbuster with that of a user-generated audio note recorded on a whim with a smartphone. In this case, though, we took a different approach and had a single production vision of creating an equivalent place-based media frame. We had controls in place, such as the same-sized budget, with the same professional expertise involved and same aspirations for the final pieces. This media covered the same event at the same time and was shared with the same general audience in the exact same place. While some differences between the forms can never be exact duplications, such as the imagery of The Old Apple Tree versus the verbal description of The Old Apple Tree, we purposively avoided those clear content concerns in favour of media composed in gray areas of the medium. Does a video of an archaeologist talking about the context and significance of The Old Apple Tree create a substantially different experience for the audience than an audio version of the same information? Visual cues, even subtle ones, certainly could provide some differences, and a close cognitive focus on sound cues could provide others. But does that matter conceptually on a medium level, suggesting one medium has superiority over another? Or are most perceived differences primarily based on artistic craftsmanship within the mediums? In this attempt to level the field and compare and evaluate media forms in frames, our data suggests that video on mobile potentially could be just as effective as an engagement tool as audio if designed and delivered in deliberate and mobile-wary ways. As such, designers and users should explore choices, not delegated by determinant dictums or social navigations, and to develop more research into this topic, which could shape the informal learning models of the future.

7   Limitations

As with any study, limitations arose. For the first phase of research, we collected a relatively small sample. During the course of the five-hour annual festival, the two researchers gathered more than ten responses an hour, and more than 50 responses total, but more responses, and more data, might have generated a deeper look at the phenomena at play. Although the study features a high quality sample, reflecting a general population in a typical state of activity, its relatively small size allows for the possibility of type II error. This does not appear to be fatal to the results for several reasons. First, none of the data came close to approaching statistical significance, which indicates a strong lack of differences between the different conditions. Further, if anything can be interpreted from

those results, sound had lower means on most variables, which further enhances the conclusions drawn from this study. And most importantly, when the data was examined in the second study, a difference still did not emerge.

As a side note, in the interest of creating the cleanest experience for the users as possible, the survey size was kept to a minimum. Participants did not receive compensation and volunteered their time on a day when they visited the site to participate in a larger festival. In that vein, all of the variables used multi-item measures, but we limited the focus of what we could investigate out of consideration for our participants' time and commitment. Although we focused on two aspects of engagement relevant to our app, O'Brien and Toms (2008) noted the complexity of this variable and how other aspects of engagement might need to be measured for a fuller understanding of the factors involved and effects. Engagement might not be the right focus for media forms in frames, after all, which might influence other elements in the environment to a more significant level. As such, future studies should continue to explore the nuances of engagement in this context. They also should further explore the intersections of the media and the location in informal learning situations.

Unlike a variety of other studies – especially those that focus upon m-Learning, and the potential of mobile technologies to enhance knowledge creation, retention, sharing, etc. – this study did not track or measure learning outcomes. Therefore, we do not know from this work what lasting impacts a medium within the mobile medium might have on, for example, information retention or processing efficiencies. In short, many aspects of this medium-focused dynamic deserve exploration, and further studies can expand the present study's scope. This study represents an initial probe and re-examination beneath the surface of this fertile ground for inquiry.

8   Implications

On a theoretical level, and in order to continue to exist as a niche medium, audio tours must demonstrate some sort of differentiation, from the perspective of the audience, in regard to the resources it provides, or through the time and space it provides access. From our research perspective, that differentiation does not appear evident in this study. If a medium cannot differentiate itself through some form of competitive superiority, it will not be able to survive, coexist, and compete for an audience, in a sort of media Darwinism outlined by Dimmick et al. (2011). Therefore, comparisons like this study illustrate that an examination of media forms operating in the same resource space, and by us manipulating only the medium frame, could help to determine, at least in the long-term, which media forms will outcompete the others on the mobile platform.

These formative findings meanwhile raise interesting questions about the roles and influences of technological determinism (Smith and Marx, 1994). The results undermine deterministic thoughts in the sense that content carefully crafted to be similar in informational value and tone – also equally situated and of the same professional quality, and delivered through a convergent digital device – appears to create generally consistent responses among its audiences, regardless of the medium within the medium. This suggests that if one had the ability to create a similarly comparable television program and radio program in the past, using the same basic tenants of our study, such as giving those to the audience in the same place and context those members would find equally meaningful, that they might have fared about the same. When they otherwise don't is

when video programming does what it does best, such as showing an action sequence, or conversely for acoustic media, provoking a listener to conjure a non-visual thought. Determinists could argue in return that without convergent media afforded through mobile devices that allow such close comparisons, video could never even compete on a practical level with audio at a site-based attraction, especially outdoors, in settings historical, such as Gettysburg National Military Park, cultural, such as the African Burial Ground National Monument, and natural, such as Yellowstone National Park. Yet with the advent of smartphones and tablet computers, such a debate and discussion about informal learning now can happen, only as determined by the development of these technologies.

Another matter to reconsider is the impact of thoughtful and carefully crafted content on its audience, regardless of the medium. In this study, we mindfully made many subtle choices about the content in both audio and video that might not have been considered by a designer who was simply trying to work well within a single medium. Those choices involved some compositional compromises, such as avoiding action sequences in the video, to mitigate advantages one medium might inherently have over the other, and to craft pieces of roughly equal situational strength. While we did our best to try to even the comparison to a level field, the larger methodological question remains: Is that even possible? Can audio ever be compared to video, even if it comes in the same small container, at the same time in the same place. And if not this way, then how? If such comparisons can be made – and we think they can, as demonstrated here – what findings can be teased from these comparisons and for what purposes? The purpose here was to challenge conventional wisdom about media choices, in a mobile-oriented and empirically oriented way, and to see what could come from it.

Hight (personal communication, March 5, 2013) noted his research team chose to produce mobile narratives in audio-only formats because the platform

> "has the power of one being fully immersed in the physical space while still engaging other layers of data/narrative. This creates a sense of dual immersion that visual data cannot, as (video) instantly inserts this hierarchical sense of here is what I am watching while I stand in some place. In locative gaming / cinema, it makes sense to chose the overlay / interplay, but in a locative narrative (especially way back in 2001), it made more sense to pass on the video."

Hight also recalled that he took the tools available at the time, including VCRs, beepers, and early cell phones, and used them the way he wanted to create the spatial narrative '34 north 118 west'. That project reimagined what the technology of the day could do to interpret a place, just like mobile projects today, like the Old Apple Tree app, continue to push our understandings of not only what can be done but also why it should be done that way, with what data, through what technologies, and in what communication frames.

This study attempted to push the research inquiry beyond artistic preferences and traditions based in cultural, historical, or technological precedents. In short, we wanted to know what the empirical evidence would show if we could compare in a systematic way two media frames interchangeable within converged media devices (the smartphone and tablet computer). In the meantime, this study undertook a novel approach to researching issues of the mobile medium, and of media frames, by building data-gathering instruments as mobile apps and deploying them into authentic situations, involving real people in real settings performing real activities. We used an authentic sample captured from the general sample, and examined both behaviour and self-reported data as part of

the analysis. On the most basic level, this study demonstrated that audio no longer should automatically deserve favour in situations that involve place-based informal learning. Medium choice, within the mobile medium, should be questioned. While audio tours have a long history, and industry support, other types of media frames might be just as valuable, or even more valuable, depending on the situation. The only way to find out is to continue to test different types of media frames in different situations with the most advanced learning technologies available, to challenge assumptions based on anecdotal evidence rather than empirical findings and to continue to push heritage communicators to produce more advanced and effective site-based informal learning materials and see – and listen to – what happens next.

References

Beer, V. (1987) 'Great expectations: do museums know what visitors are doing?', *Curator: The Museum Journal*, Vol. 30, No. 3, pp.206–215.

Bolter, J. and Grusin, R. (1998) *Remediation: Understanding New Media*, MIT Press, Cambridge, MA.

Crompton, H. (2014) *A Diachronic Overview of Technology Contributing to Mobile Learning: A Shift Towards Student-Centred Pedagogies Increasing Access Through Mobile Learning*, pp.7–15, Commonwealth of Learning and Athabasca University, Vancouver, BC.

Dimmick, J., Feaster, J.C. and Hoplamazian G.J. (2011) 'News in the interstices: the niches of mobile media in space and time', *New Media & Society*, Vol. 13, No. 1, pp.23–39.

Dow, S., Lee, J. Oezbek, C., MacIntyre, B., Bolter, J.D. and Gandy, M. (2005) 'Exploring spatial narratives and mixed reality experiences in Oakland cemetery', *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology*, ACM, Valencia, Spain.

Druckman, J.N. (2001) 'The implications of framing effects for citizen competence', *Political Behavior*, Vol. 23, No. 3, pp.225–256.

Entman, R.M. (1993) 'Framing: toward clarification of a fractured paradigm', *Journal of Communication*, Vol. 43, No. 4, pp.51–58.

Fagerjord, A. (2011) 'Between place and interface: designing situated sound for the iPhone', *Computers and Composition*, Vol. 28, No. 3, pp.255–263.

Falk, J. and Dierking, L. (2008) 'Enhancing visitor interaction and learning with mobile technologies', in Tallon, L. (Ed.): *Digital Technologies and the Museum Experience: Handheld Guides and Other Media*, pp.19–34, AltaMira Press, Lanham, MD.

Gitlin, T. (1980) *The Whole World is Watching: Mass Media in the Making & Unmaking of the New Left*, University of California Press, Berkeley, CA.

Hight, J. (2005) 'Narrative archaeology: reading the landscape', *Proceedings of MiT4: The Work of Stories*, Cambridge, MA.

Hight, J. (2006) 'Views from above: locative narrative and the landscape', *Leonardo Electronic Almanac*, Vol. 14, Nos. 7/8, pp.1–9.

Hjorth, L. (2013) 'Relocating the mobile: a case study of locative media in Seoul, South Korea', *Convergence: The International Journal of Research into New Media Technologies*, Vol. 19, No. 2, pp.237–249.

Höök, K. and Löwgren, J. (2012) 'Strong concepts: intermediate-level knowledge in interaction design research', *ACM Transactions on Computer-Human Interaction (TOCHI)*, Vol. 19, No. 3, p.23.

Jacobs, R.D. (1999) 'Evaluating customer satisfaction with media products and services: an attribute-based approach', *European Media Management Review*, Vol. 32, No. 1, pp.11–18.

Jasim, W.H. and AbuSeileek, A.F. (2015) 'Using text and voice chat modes to enhance students' performance in discourse functions', *International Journal of Learning Technology*, Vol. 10, No. 2, pp.114–136.

Jenkins, H. (2006) *Convergence Culture: Where Old and New Media Collide*, NYU Press, New York.

Kang, M. and Gretzel, U. (2012) 'Perceptions of museum podcast tours: effects of consumer innovativeness, internet familiarity and podcasting affinity on performance expectancies', *Tourism Management Perspectives*, Vol. 4, No. 1, pp.155–163.

Kay, A. and Goldberg, A. (1977) 'Personal dynamic media', *Computer*, Vol. 10, No. 3, pp.31–41.

Klein, G., Moon, B. and Hoffman, R.R. (2006a) 'Making sense of sensemaking 1: alternative perspectives', *IEEE Intelligent Systems*, Vol. 21, No. 4, pp.70–73.

Klein, G., Moon, B. and Hoffman, R.R. (2006b) 'Making sense of sensemaking 2: a macrocognitive model', *Intelligent Systems*, IEEE, Vol. 21, No. 5, pp.88–92.

Klein, G., Phillips, J.K., Rall, E.L. and Peluso, D.A. (2007) 'A data-frame theory of sensemaking', in Hoffman, R.R. (Ed.): *Expertise out of context: Proceedings of the Sixth International Conference on Naturalistic Decision Making*, pp.113–155, Lawrence Erlbaum Associates, New York, NY.

Maniar, N., Bennett, E., Hand, S. and Allan, G. (2008) 'The effect of mobile phone screen size on video based learning', *Journal of Software*, Vol. 3, No. 4, pp.51–61.

Mersey, R.D., Malthouse, E.C. and Calder, B.J. (2010) 'Engagement with online media', *Journal of Media Business Studies*, Vol. 7, No. 2, pp.39–56.

Miranda, M. (2008) 'Directions in locative media', *Multimedia*, IEEE, Vol. 15, No. 3, pp.4–5.

Novey, L. and Hall, T. (2007) 'The effect of audio tours on learning and social interaction: an evaluation at Carlsbad Caverns National Park', *Science Education*, Vol. 91, No. 2, pp.260–277.

O'Brien, H.L. and Toms, E.G. (2008) 'What is user engagement? A conceptual framework for defining user engagement with technology', *Journal of the American Society for Information Science & Technology*, Vol. 59, No. 6, pp.938–955.

O'Brien, H.L. and Toms, E.G. (2010) 'The development and evaluation of a survey to measure user engagement', *Journal of the American Society for Information Science & Technology*, Vol. 61, No. 1, pp.50–69.

Patwardhan, P. Yang, J. and Patwardhan, H. (2011) 'Understanding media satisfaction: development and validation of an affect-based scale', *Atlantic Journal of Communication*, Vol. 19, No. 3, pp.169–188.

Peart, B. (1984) 'Impact of exhibit type on knowledge gain, attitudes, and behavior', *Curator: The Museum Journal*, Vol. 27, No. 3, pp.220–237.

Piper, A.M., Friedman, W. and Hollan, J.D. (2012) 'Setting the stage for embodied activity: scientific discussion around a multitouch tabletop display', *International Journal of Learning Technology*, Vol. 7, No. 1, pp.58–78.

Pirolli, P. and Russell, D.M. (2011) 'Introduction to this special issue on sensemaking', *Human-Computer Interaction*, Vol. 26, Nos. 1–2, pp.1–8.

Richardson, I. (2011) 'The hybrid ontology of mobile gaming', *Convergence: The International Journal of Research into New Media Technologies*, Vol. 17, No. 4, pp.419–430.

Samis, P. (2008) 'The exploded museum', in Tallon, L. (Ed.): *Digital Technologies and the Museum Experience: Handheld Guides and other Media*, pp.3–18, AltaMira Press, Lanham, MD.

Smith, M.R. and Marx, L. (1994) *Does Technology Drive History?: The Dilemma of Technological Determinism*, MIT Press, Cambridge, MA.

Soo, L.K., Yeoh, E-T. and Ho, S-B. (2014) 'Personalised reusability metric for learning objects', *International Journal of Learning Technology*, Vol. 9, No. 4, pp.356–373.

Stolterman, E. and Wiberg, M. (2010) 'Concept-driven interaction design research', *Human-Computer Interaction*, Vol. 25, No. 2, pp.95–118.

Suh, Y., Shin, C., Woo, W., Dow, S. and MacIntyre, B. (2011) 'Enhancing and evaluating users' social experience with a mobile phone guide applied to cultural heritage', *Personal and Ubiquitous Computing*, Vol. 15, No. 6, pp.649–665.

Tallon, L. (2008) *Digital Technologies and the Museum Experience: Handheld Guides and Other Media*, Altamira Press, Landham, MD.

Taylor, J., Sharples, M., O'Malley, C., Vavoula, G. and Waycott, J. (2006) 'Towards a task model for mobile learning: a dialectical approach', *International Journal of Learning Technology*, Vol. 2, Nos. 2–3, pp.138–158.