Mediating Attention for Second Screen Companion Content

Timothy Neate
FIT Lab, Swansea University
Swansea, SA2 8PP UK
tdjneate@gmail.com

Matt Jones
FIT Lab, Swansea University
Swansea, SA2 8PP UK
mattjonez@gmail.com

Michael Evans
BBC R&D
Salford, M50 2LH UK
michael.evans@rd.bbc.co.uk

ABSTRACT

There is increasing interest in providing content to users on secondary devices while they watch TV. This material, termed companion content, can be anything from textual information, to interactive quiz games. It can be delivered throughout a broadcast and often directly relates to specific scenes in a show. This new scenario has exposed a challenging design space for creators of both the content and the enabling technology. A key question when introducing content on a secondary device is how much it detracts from, or enhances, the show the user is currently engaged with. To examine this, we investigated methods for mediating attention from the TV and onto a secondary device. By examining a typical use case we have been able to gain new insights into how best to design additional stimuli to alert users to companion content from both a broadcasting, and an HCI perspective.

Author Keywords

Companion content; second screens; attention; TV; Alerts

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

Multi-screen TV experiences are becoming increasingly prevalent, whether instigated by the broadcaster or as an independent behaviour by the viewer. We might browse the web for related content on our phone or computer whilst watching TV; perhaps Googling an actor we recognise in a show when we cannot place him. However, the material we browse is not necessarily related to the content we are viewing. As described in the work of Rooksby et al. [6], this relationship between the two devices often becomes a complex web of related, semi-related, and non-related content.

A desire to leverage this dual screen interaction and enhance the user experience (UX) of TV, has led to broadcast content providers developing companion content. Regardless of the application, we believe that the design of attention across devices should be a part of the craft of TV making – primary and companion content should enhance, and not disrupt, each other. A potential innate tension in using second

screens to enhance the UX of TV is the fact that content on the secondary device can clearly have a distracting effect on the attention available for the primary content. Therefore, it is crucial to understand methods by which TV producers and engineers can design and deploy companion content without detracting from the experience as a whole. Though much research has considered visual attention for TV-based scenarios, for example the recent work of Valuch et al. [9] and Vatavu et al. [10], no work has explored the notion of how we may mediate attention from the motion-rich television to the companion content. Therefore, in this paper we explore methods for attracting a user's attention between screens from an objective and subjective standpoint so that we may begin to consider how we may orchestrate companion content.

BROADCASTING IN AN IP WORLD

From both the production and the audience perspective, the notion of what TV broadcasting *is* is undergoing significant change. The main technical driver for this is the end-to-end use of IP networks. The UX trend is heading away from scheduled programmes that are identical for everyone. Traditional linear production of TV involves the capture of signal-based assets, from cameras, microphones and other sources, together with a live, concurrent or post-production editing and compositing phase, where those media assets are used to construct a bundled programme. This piece of content is then delivered, in an essentially identical form, for equivalent consumption by each member of the audience: the programme is an immutable block.

Conversely, IP-based broadcasting – featuring data-format agnostic bi-directional communication between audience and broadcaster is not constrained in this way. Creating and broadcasting content as a related set of media and data objects, together with curatorial decisions describing its recombination, means that TV content can be made responsive to a number of human factors and use cases. This lends itself to significant new possibilities for media experiences such as those discussed in depth by Armstrong et al. [1], but more pertinently, such methods of broadcast yield considerable possibilities for second screen experiences.

The terminology of second screens and companion content refers to a spectrum of dual screen TV use cases. Essentially, companion content is second screen material that embellishes the primary TV experience, such as factual text or interactive games. Broadcasters and the HCI community have developed numerous examples of this type of experience in recent years. Two notable examples have formed some preliminary UX considerations for this design space: a companion application for tablets to support experimental studies alongside
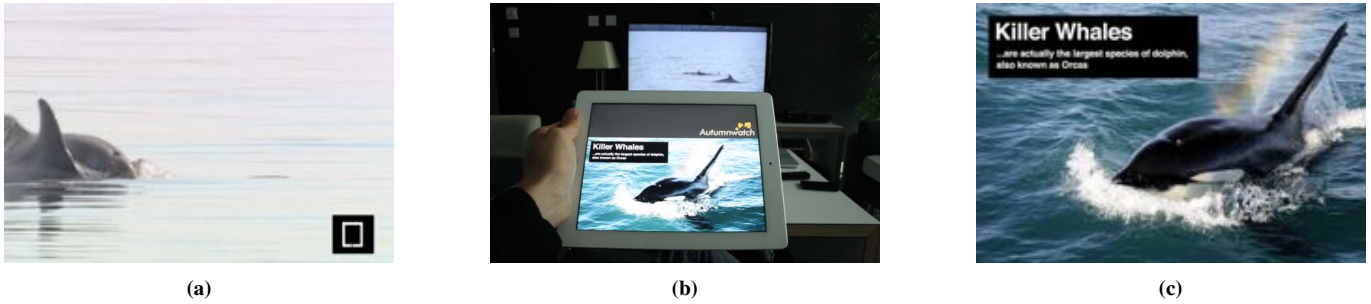
Figure 1: Examples of content and stimuli used: (a) shows an icon indicating companion content in the bottom right of the TV, (b) shows a user watching the companion content along with the show, and (c) shows this second screen content close up.

No.	Type	Description
1	Content appears	Content simply appears on tablet
2	Content shakes	Content shakes on tablet
3	Earcon	User notified with musical sound
4	Auditory icon	User notified with related sound
5	On TV	Icon appears on bottom right of TV
6	On TV & shaking	Icon shakes in bottom right of TV

Table 1: Methods used to attract the attention from the TV.

the nature programme Autumnwatch [1], and a quiz game to accompany the programme Antiques Roadshow [2].

For Autumnwatch, companion content appeared on a second screen during the TV broadcast. To explore this use case more systematically, Brown et al. [3] observed gaze across both screens while users watched this content. The findings indicate that, unlike task-based scenarios, where division of attention is divided more equally between the screens [5], users focus primarily on the TV and only intermittently pay attention to companion content. This study implied that with such a focus on the primary (TV) screen, it is critical to design for a positive second screen experience without disrupting the UX of the programme.

ATTENTION MEDIATION METHODS
The aim of this experiment was to find the best methods for getting a user's attention to transfer to the companion content in the most efficient and preferable way. To this end we developed methods and tested them in a typical scenario.

Design of Stimuli
As described in Table 1, there were 6 methods tested – one (baseline) where the content just appears on the tablet (no. 1), one animated visual method on the tablet (no. 2), two auditory methods on the tablet – an earcon (no. 3) and an auditory icon (no. 4) (both defined in [2]). Also, there were two methods that notified the participants on the TV itself by providing an icon (Figure 1a) – one stationary (no. 5), and one animated (no. 6).

We designed the shaking visual alerts on the device (no. 2) and on the TV (no. 6) to best gain attention and negate change

[1]Autumnwatch app – http://goo.gl/hI3UJs (accessed 18/6/14)
[2]Autumnwatch app – http://goo.gl/hI3UJs (accessed 18/6/14)

blindness [8]. As peripheral vision is poor, in terms of acuity, but tuned to detect motion (specifically horizontal) [4] we designed them to shake diagonally to attract a user's gaze. To avoid the content going unnoticed and combat the auditory equivalent of change blindness we designed two types of auditory notifications – one totally disjunct from the sounds in the show (no 3); and a more related type (no. 4). The earcon (no. 3) featured a musical sound, which was designed to be totally different from the sounds associated with the show (natural sounds, incidental music, and presenters speaking); and, the auditory icon (no. 4) featured related sounds – for example, the sound of a river running when a fish is in shot on the TV.

METHOD
We recruited 18 participants from general staff mailing lists at the BBC. Their ages ranged from 25 to 48, with a mean of 36 ($\sigma = 7.24$), with 11 identifying as male and 7 female. Users reported an average of 2.53 hours of TV viewing per day ($\sigma = 1.43$), 16 (89%) of the participants stated that they engaged with mobile devices while watching TV, and 10 (56%) stated that they are easily distracted by other electronic devices while watching TV.

We conducted the study in a usability lab, rigged as a generic living room. The participants were in a typical scenario for such a viewing experience – sitting on a sofa in front of a 42 inch modern TV with a tablet computer (iPad 2) on their lap. As depicted in Figure 2, participants were filmed from directly in front by a camera on top of the TV, and also from above the iPad, allowing for us to make recordings of the participants and use them for analysis. The participants were first talked through the study and then asked to fill in the consent and demographics forms. They then watched the television show, which was an excerpt from Autumnwatch – similar to content used by Brown et al. [3]. During this, second screen content was introduced at 2 to 4 minute intervals to complement the programme. This content was simple images and textual information (as in Figure 1c). Each time a piece of content appeared it was accompanied by an attention mediating method (for example the content appeared on the tablet and shook). Upon completion, the user was asked to fill in a post-study questionnaire that allowed us to gauge their impressions.
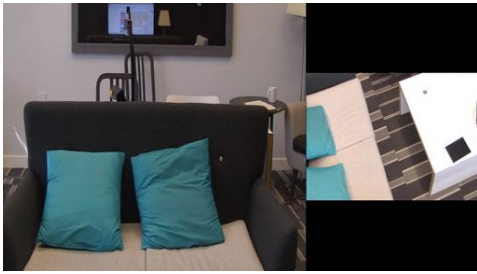
Figure 2: Camera angles used for analysis: the left side (TV mounted camera) allowed for a good view of the participant's gaze, and the right side allowed for us to determine when new content had appeared.

In terms of experimental design, we used a Latin square arrangement to ensure that the participants were exposed to the methods in a different order each time so that we could observe their effect irrespective of the content on the TV or the tablet. Reaction times were our primary objective metric. To observe these, we inspected the time between the content appearing on the tablet and each participant's attendance to it. By analysing the video (Figure 2) we could track the participants' eyes and easily infer where they were looking and at exactly what time. To ensure consistency we adhered to a strict set of criteria to classify the participants' gaze, and verified our results with a second sampled analysis, and with an external party. We then framed our objective findings by using a post-study questionnaire to gauge how much the methods got their attention and to assess what they preferred.

RESULTS
In this section we describe the objective findings (reaction times). We then discuss the participants' subjective impressions of the stimuli by analysing their ratings, and briefly discuss their comments.

Objective Analysis
Figure 3 shows that peripheral methods (on-device auditory and visual stimuli) resulted in the fastest overall reaction times from the participants, and that notifying the participants on the television resulted in the slowest. We conducted ANOVA to examine the data and found significant overall variance ($F_{(6,17)} = 2.826$, $p = .037$). Moreover, upon conducting post-hoc tests on the stimuli conditions, we find significant differences between the Earcon and the stimuli appearing on the television, both static ($p = .001$), and shaking ($p = .020$), and also compared to the content appearing on the device ($p = .041$). Notably, a shaking alert on the iPad appears to attract attention to the tablet quicker than the non-animated cue, but the opposite occurs on the TV.

Participants' Impressions
Participants were asked to describe how well they believed different cues had performed in attracting their attention to the tablet. As shown in Table 2, 11 participants (62%) strongly agreed that the earcon was most effective at attracting their attention, followed by the auditory icon. Both were preferred to the content simply appearing without additional stimuli – significantly so for the earcon ($Z = 2.59$, $p = .005$).
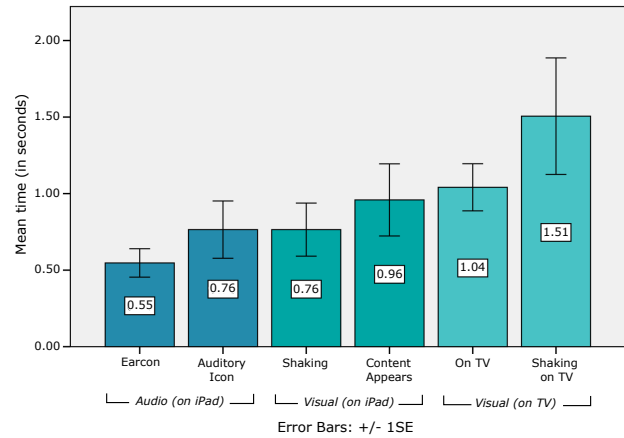

Figure 3: Participants' reaction times in ascending order.

In terms of subjective preferences, auditory icons (the sounds that related directly to the show) were most favoured, significantly more so than the content simply appearing ($Z = 2.03$, $p = .021$). Also, approval of both TV-based alerts was significantly higher than the no alert case ($Z = 2.16$, $p = .016$ and $Z = 1.86$, $p = .032$ for non-shaking and shaking respectively). Several participants noted that they used the notification on the TV to delay when they viewed second screen content, for example – "*I think the TV icon was best, I can look at the 2nd device in my own time then*".

	Missed	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Content appears	2	0	2	1	7	6
Content shakes	6	1	0	0	4	7
Earcon	4	0	0	0	3	11
Auditory Icon	3	1	0	0	5	9
On TV	3	0	3	2	3	7
Shaking on TV	4	0	1	2	3	8

Table 2: Participants' reported attention gained for each method. Darker colours indicate higher frequencies.

The uncued (content appears) case scored very poorly, and comments from the participants imply that they believed that they would miss the content if not alerted to it – "*[I] was constantly checking to see if I'd missed it.*" – "*It felt a little surprising just to discover content*". On the other hand, the highest scoring method, the auditory icon, was mostly praised because of its ability to link the show to the second screen content, for example – "*It felt [a] more sympathetic and complimentary way of identifying that more content was available.*" and "*it felt less intrusive*". Conversely, it was apparent that some participants found audio distracting in general and would have preferred more discreet methods such as vibration. Notably, we found a correlation between how much a user believed something got their attention, and how much they liked it (Spearman's $\rho = .384$, $p < .001$).

Table 3 shows that a notable proportion (39%) did not notice the notification shake on the TV. Some comments indicated that they did not notice such subtleties because their attention

	Missed	Strongly Disagree	Disagree	Neutral	Agree	Strongly agree
Content appears	2	1	6	3	4	2
Content shakes	3	0	0	9	5	1
Earcon	1	0	4	4	7	2
Auditory Icon	2	0	1	5	6	4
On TV	4	0	1	4	7	2
Shaking on TV	7	0	1	4	2	4

Table 3: Participants' reported preference for each method. Darker colours indicate higher frequencies.

had already been diverted to the tablet – "*because it happened at the same time [as the alert on the tablet] I felt unprepared. The [secondary] device caught my attention*".

DISCUSSION

Reaction time data and the subjective feedback indicate that auditory methods performed best at quickly attracting users' attention to companion content. Likely causes of this observation are: a) we process auditory information significantly faster than visual [7], b) that peripheral stimuli (in general) cause a strong disjunct from the TV show, and c) that, in the case of the earcon, its lack of ecological coherence with the TV content forms a more explicit highlighting of change in the content on the tablet. Note, however, that reaction time may not be the most important measure: the subjective data shows that not only do participants' opinions vary greatly, but also that the fastest method (the earcon) for getting attention onto the tablet is not generally as favoured as the auditory icon.

We believe there are two reasonable explanations for the notifications on the TV being the slowest: Firstly, several participants reported that they used it as a method of delaying a voluntary switch of attention. Secondly, main screen alerts attracted (and split) participants' attention first delaying a shift of attention to the second screen. Moreover, we believe that the shaking effect on the TV did not aid the participants in noticing the content on the iPad quicker was because its motion focused their gaze onto the TV and away from the tablet in their periphery.

CONCLUSIONS AND DESIGN IMPLICATIONS

This paper investigates methods for attracting attention to companion content when a user is engaged with the related TV programme. We found that users react quicker to the availability of refreshed second screen content when alerted by means of peripheral stimuli – especially through the modality of sound. In addition, users much prefer their attention to be mediated, to insure against missing parts of primary content (constantly glancing at the tablet to check for new content), or second screen content (missing it because they are too focused on the TV).

Interestingly, our findings indicated that many participants liked alerts on the TV itself, as it supported consciously delaying shifting their attention to the companion content until a moment of their choosing. Designers of dual screen user experiences should consider how closely in time the companion content needs be synchronised to the main screen content and cue the user appropriately. If, for example, the content is relevant to only a few seconds of the show, a designer may wish to use attention mediating techniques which have been found to cause more involuntary, faster attendance to the second screen, such as peripheral/auditory methods. However, if the relevance of the content to the notification cues spans minutes rather than seconds, then perhaps methods that use notifications on the main TV screen may be more beneficial, as these will allow users time to choose when to look at them.

REFERENCES
1. Armstrong, M., Brooks, M., Churnside, A., Evans, M., Melchior, F., and Shotton, M. Object-based broadcasting: Curation, responsiveness and user experience. In *Proc. Int. Broadcasting Convention* (2014), 12.2.

2. Blattner, M., Sumikawa, D., and Greenberg, R. Earcons and icons: Their structure and common design principals. *Human-Computer Interaction 4* (1989), 11–44.

3. Brown, A., Glancy, M., Evans, M., Jones, R., Jay, C., and Harper, S. Hci over multiple screens. In *CHI EA '14* (2014), 665–674.

4. Fahle, M., and Wehrhahn, C. Motion perception in the peripheral visual field. *Graefe's Archive Ophthalmology*, 229 (1991), 430–436.

5. Grudin, J. Partitioning digital worlds: Focal and peripheral awareness in multiple monitor use. In *Proc. CHI*, ACM (2001), 458–465.

6. Rooksby, J., Rost, M., Morrison, A., Bell, M., Chalmers, M., and Smith, T. Practices of parallel media: Using mobile devices when watching television. In *Workshop at CSCW – Designing with Users for Domestic Environments: Methods, Challenges and Lessons Learned* (Feb. 2014).

7. Shelton, J., and Kumar, G. P. Comparison between auditory and visual simple reaction times. *Neuroscience and Medicine 1* (Sept. 2010), 30–32.

8. Simons, D., and Levin, D. Change blindness. *Trends in Cognitive Sciences 7*, 1 (1997), 261–267.

9. Valuch, C., Ansorge, U., Buchinger, S., Patrone, A. R., and Scherzer, O. The effect of cinematic cuts on human attention. In *Proc. TVX*, ACM (2014).

10. Vatavu, R.-D., and Mancaş, M. Visual attention measures for multi-screen tv. In *Proc. TVX*, ACM (2014).