Investigating the Effect of Relative Time Delay on Companion Screen Experiences

Chua Wei Liang Kenny[1], Jacob M. Rigby[1], Duncan P. Brumby[1], Vinoba Vinayagamoorthy[2]
[1]UCL Interaction Centre, University College London, London WC1E 6EA UK
[2]BBC R&D, BBC Centre House, London W12 7SB UK
[1]{wei.chua.15, j.rigby.14, d.brumby}@ucl.ac.uk
[2]vinoba.vinayagamoorthy@bbc.co.uk

ABSTRACT

Mobile devices are increasingly used while watching television, leading to the development of companion apps that complement the content of programmes. A concern for these applications is the extent to which companion app and television content need to be temporally aligned for live synchronisation. In this study, 18 participants watched a nature programme while being shown companion content on a tablet. Temporal synchronisation of content between the devices was varied. Participants completed questionnaires measuring immersion and affect and were tested on their recall for companion app content. While there were no statistically significant effects on these measures, qualitative interviews with participants after viewing consistently revealed that longer 10s delays in content synchronisation were frustrating. This suggests that poor content synchronisation can produce a negative companion experience for viewers and should be avoided.

Author Keywords

Companion Screen; Connected TV; Mobile; Multilevel Measurements; Multiscreen Interaction; Second Screen; Media Multitasking; Smart Phones; Synchronisation

ACM Classification Keywords

H.5.2 [Information Interfaces and Presentation]: User Interfaces – User-Centered Design.

INTRODUCTION

The rapid uptake of powerful mobile devices has led to concurrent device usage and television viewing becoming common [20]. This "media multitasking" allows viewers to perform other activities while viewing, e.g. researching an actor, or exchanging opinions on social media.

Some content producers have developed specially designed companion apps to be used during a programme to enrich or augment viewing. These applications can rely on good synchronisation to deliver time-sensitive content to the user, which can be difficult to achieve at present due to limitations in synchronisation technology [24].

The study presented here investigates how viewing time-sensitive companion screen experiences with differing levels of relative time delay affects immersion, affect, and content recall. We first introduce work relevant to understanding companion screen experiences and methods of measuring engagement. We then describe our study and present the results. We end by discussing how our results can inform the development of multi-screen experiences and suggest directions for future work.

RELATED WORK

TV companion app experiences

Ad-hoc companion experiences have been examined previously in the literature. These have included systems that increase accessibility to disabled users [24], and add synchronized companion content to the main program [1]. Users reacted positively to the availability of additional information without needing to look it up themselves.

A critical issue for companion apps is that content often has a strong temporal relevance. It is possible that introducing content outside of this window of relevance could result in confusion or frustration for the viewer. For example, Geerts et al. [8] observed households that used a companion app linked to a television drama. The companion app delivered content at the same time as relevant events on the main television screen (e.g., a text message sent to a character in the show could be read on the viewer's mobile device). During interviews, participants said that they enjoyed this media experience when the companion app content was well synchronized with the show. However, participants were easily annoyed if content was delivered late because of a lack of synchronisation between devices. Geerts et al. [8] also interviewed content producers, who stressed that tight synchronisation was difficult to achieve with the technology available at the time.

A number of approaches are currently used to achieve synchronisation, such as audio watermarking or fingerprinting [24], IP-based approaches, and direct

communication between the TV and the companion device [2, 10, 25]. Until these technologies are commonplace and tight synchronisation can be achieved, temporal delays will be present and the effects should be investigated.

Measuring engagement and viewer experience
Despite the importance of understanding viewer engagement, there is still no clear definition or means of measuring engagement and viewer experience [16]. This has given rise to a variety of ways to measure TV engagement, such as a questionnaire to measure narrative engagement [9] and tracking of activities on social networking site regarding the show content [18]. Other subjective measures include a measure of arousal and valence collected through questionnaires, such as the Positive and Negative Affect Schedule (PANAS) [29].

In contrast, video game researchers have had success with using *immersion* as a way of gauging a player's engagement with the use of the Immersive Experience Questionnaire (IEQ) [12]. Recent studies have also explored the use of a modified IEQ to measure the engagement levels of TV viewers when using different screen sizes [22] and for working while watching TV [4]. The modified Film IEQ kept the main structure of the original IEQ [12] but made changes to selected questions based on adaptations from the narrative engagement questionnaire [9]. This was done to consider the differences between watching film and playing games.

Physiological measures can complement measuring engagement via questionnaires. This has been useful to assess engagement in advertising media [27], e-learning platforms [28] and responses to media degradations [30, 31]. Two widely used dimensions for describing emotions are *arousal* (activation/excitement levels) and *valence* (positive or negative aspect of the emotion) [14], which are sensitive to short-term changes which can otherwise be difficult to detect. *Galvanic skin response* (GSR) and *heart rate* (HR) are commonly used measures of arousal and valence respectively [26, 30]. GSR measures changes in skin conductance, and has been closely linked to measuring arousal and attention [30]. HR is an indicator of overall activity level, with a higher HR representing a more anxious state. Advances in wearable technology has made collecting this data trivial and non-invasive [7].

CURRENT STUDY
We conducted a study where participants watched a nature programme while being presented with companion content on a tablet. Temporal synchronisation of content between the devices was systematically varied. In a baseline condition, relevant content appeared on the tablet when it was referred to in the programme. Further conditions introduced delays of either 3s or 10s to content appearing after being referred to in the programme. We might expect that longer time delays will have a higher negative effect on participants' responses to companion screen experiences (i.e., lower self-reported affect and immersion scores).

We wanted to study whether time delays in the delivery of content to a secondary screen affects when viewers choose to look at that screen, and if this has implications for missing TV content. We might expect to find that longer time delays cause a greater interruption effect, leading to participants switching attention between screens more often and missing companion content.

Finally, we attempted a preliminary investigation into participants' physiological responses to time delays in companion screen experiences. We expect physiological measures should correlate with self-reports of affect.

METHOD

Participants
Participants were screened to ensure they had not seen the BBC nature documentary *Alaska: Earth's Frozen Kingdom*, which was used as part of the stimulus. There were 18 participants (8 female), aged 22-39 (*M*=28.1, *SD*=4.9), and all were native English speakers with normal or corrected vision. 67% of the participants consumed at least 10 hours of video content per week, and 78% used a companion device while watching. Participants had a chance of winning Amazon vouchers for an hour of their time.

Design
A one independent variable (time delay) counterbalanced within-subject design was used. Three levels of relative time delay (0, 3s, 10s) values were chosen after making minor adjustments to the guidelines provided by Nielsen [19]. The 0s condition (no injected relative time delay) was a baseline condition. As it takes at least 800ms to 2s [3, 11, 17] for people to glance at companion device, we choose a time delay value of 3s for the mid-level value. The third test value was 10s, as people tend to lose attention after that point [15]. These values were within the realistic time delay ranges in current synchronisation technology [24].

We were primarily interested in the effect of injected delays between the TV and the tablet, while presenting a supposedly synchronised companion screen experience. We used a multi-level approach to assess this: the Film IEQ [22] and PANAS [29] questionnaires were used to measure immersion and affect respectively; an Empatica E4 wristband [7] measured the GSR and HR as additional indicators for arousal and valence; a companion content recall test was conducted at the end of the study to see if participants could recall what they had seen; and semi-structured interviews were used to see if participants sensed a time delay, judged a time delay estimation of each condition and to gather general feedback.

Materials
Three similar 7-minute clips from the documentary "*Alaska: Earth's Frozen Kingdom*" were shown on the TV. 21 images from the BBC database and wider Internet were used to create seven distinct one-minute slides to be used as companion content for each of the three clips, each containing an image and 15-20 words. The slides were of

similar visual complexity (lowest visual complexity metric of the Flesch [13] reading ease score: 90-100).

Each slide had an exposure time of 15s, based on an average reading rate of 200 words per minute [21]. After appearing on the tablet for 15s, the content disappeared leaving the background scene. A visual notification appeared on the bottom right of the TV informing the participants that companion content was available, following a methodology used in a previous study [17].

In the baseline condition, companion content appeared on the tablet in sync with the programme timeline on the TV as designed. In the relative delay conditions, companion content appeared on the tablet in sync with the programme timeline on the TV for the first two minutes of the 7-minute experience. This allowed us to collected rest baseline physiological readings of the participants. For the remaining 5 minutes the notification icon appeared at the correct times, but a black screen with a loading icon was delayed for the relevant time delay (3s or 10s) before transitioning to the companion content.

Participants sat on a sofa in a usability lab, in front of a 30-inch monitor connected to a laptop simulating a connected TV. Participants used an iPad 2 tablet running a prototype companion screen application. The TV and tablet were connected to the same Wi-Fi network. The prototypes on both the TV and the tablet had built in DVB-CSS protocols which enabled the synchronised playback of video/audio files and presentation of web content, to frame-level accuracy, between the devices [2, 10, 25]. An Empatica E4 wristband was used to collect physiological data [7].

Procedure
Participants were greeted, briefed, and asked to sign a consent form. They were then asked to wear an E4 wristband and allowed 5 minutes to get comfortable to collect a resting baseline reading [14]. They were asked to fill in a questionnaire to collect demographic data. The participants were then told that the experiment was designed to study user behaviour, and that they would watch three clips from a factual documentary. They were advised to hold the companion screen on their lap to act as a control but were free to adjust the position if desired. A sample companion screen experience was shown as a demo. Participants were left alone in the room for each condition. After each condition, the experimenter returned to the room to administer questionnaires. After completing all the conditions, participants were administered a companion content recall test followed by a semi-structured interview.

RESULTS

Self-reported Immersion
Immersion scores for each condition were computed as in [12, 22]. Participants had higher mean immersion scores in the 10s delay condition ($M = 154.55$, $SD = 23.02$) than in the 0s baseline condition ($M = 149.72$, $SD = 18.30$) and the 3s delay condition ($M = 149.83$, $SD = 23.31$). However, a one-way repeated measures Friedman's ANOVA found no significant effect of relative time delay on immersion scores. Friedman's ANOVA tests were performed on each Film IEQ subscale, but found no significant effects.

Negative and Positive Affect Scores
Positive and negative affect score were computed for the PANAS questionnaires [29] administered to each participant. As shown in Table 1, mean scores were similar across conditions. Indeed, a one-way repeated measures Friedman ANOVAs found that there was no significant effect of relative time delay condition on PANAS scores. Further investigation of the PANAS subscales, using a series of one-way repeated measures Friedman's ANOVA, revealed that participants were significantly more upset in the 10s time delay condition in comparison to the other two conditions, $X^2 (2) = 12.5$, $p < 0.05$.

Time Delay	Positive Affect (Mean ± Std. Dev.)	Negative Affect (Mean ± Std. Dev.)
0s	26.28 ± 5.60	13.22 ± 3.37
3s	26.33 ± 6.62	12.17 ± 2.66
10s	25.28 ± 5.85	14.28 ± 3.03

Table 1. PANAS Scores by Delay Condition.

Galvanic Skin Response and Heart Rate
To analyse participants' tonic response, we measured the slope of the GSR signal during the time that the delayed companion content was shown. Participants had a smaller change in GSR in the 0s baseline condition ($M = 0.29\mu S$, $SD = 0.62\mu S$) than in the 3s relative delay condition ($M = 0.35\mu S$, $SD = 0.68\mu S$) or in the 10s relative delay condition ($M = 0.34\mu S$, $SD = 1.46\mu S$). However, no significant effect of relative time delay on GSR data was found.

We computed mean HR using the same time window used for the GSR responses. Participants had a lower HR in the 0s baseline condition ($M = 71.21$ bpm, $SD = 12.48$ bpm) than in the 3s relative delay condition ($M = 72.23$ bpm, $SD = 12.74$ bpm) or in the 10s relative delay condition ($M = 72.06$ bpm, $SD = 12.58$ bpm). However, no significant effect of relative time delay on HR data was found.

Time Delay	Hit Rate, False Alarm Rate (%, %)	Memory Strength (Mean ± Std. Dev.)
0s	87%, 1%	6.58 ± 2.27
3s	84%, 2%	6.17 ± 2.22
10s	79.1%, 4%	5.30 ± 2.60

Table 2. Memory Recall Measures by Delay Condition.

Companion Content Recall Score
In this test, participants responded "yes" or "no" when shown a series of 30 slides, 15 of which had not been seen by the participant. The hit and false alarm rates were then used to assess the participants' memory, as in [5]. Table 2

shows that longer time delay conditions were causing people to recall less companion content. However, a one-way repeated measures Friedman's ANOVA showed no significant differences between conditions.

Relative Time Delay Estimation Error

All participants identified the 10s relative delay condition, and only a handful of them were unable to sense a difference between the 3s relative delay and the baseline condition. We computed the difference between estimated time delay value and ground truth [6] (see Table 3). A one-way repeated measures ANOVA showed that time delay estimation was significantly affected by the different levels of relative time delays, $F(2, 34) = 149.83$, $p<.001$. A series of post-hoc t-tests were performed to examine pairwise differences between each condition, using Bonferroni corrections. Results indicate a significant difference in relative time delay estimation error between all conditions; baseline and 3s time delay, $t(17) = -4.12$, $p<.001$, baseline and 10s time delay, $t(17) = -13.77$, $p<.001$, and 3s and 10s time delays, $t(17) = -4.16$, $p<.001$.

Time Delay	Mean and Std. Dev.
0s (baseline)	0.31 ± 0.46
3s	1.44 ± 1.03
10s	6.42 ± 1.70

Table 3. Time Delay Estimation Error by Delay Condition.

Interview Findings

Eighteen post-experiment interviews were coded using thematic analysis. This gave 83 references, which were grouped into key themes, including reasons for using/not using the companion application, strategies to mitigate time delays, opinions on the TV notification icon, and suggestions for improvement.

Eight participants said they would use the companion application again, most commonly because the companion content provided relevant information (7 of 18). Participants also liked not having to search for information online (2 of 18), and the "filling in" during quieter parts of the programme (7 of 18). These are like the findings of previous research [1, 8, 24]. The remaining 10 participants, however, said they would not use it again, citing their preference to focus on the TV (4 of 18), and their dislike of having to switch to another device (5 of 18). Three participants said they felt more anxious when the companion application was in use due having to monitor it for updates. Some participants stated they might use it for other genres (4 of 18) or if there was interactivity (3 of 18) which let them search for information at their own pace.

Fourteen participants said the 10s time delay condition was annoying, and that synchronised content reduced the need to constantly monitor the tablet. Three participants thought the 10s delay was a technical glitch. Interestingly, participants talked of devising different strategies to adapt to the delay. This included ignoring the companion screen notification on the TV, and relying on their peripheral vision to detect the update directly from the tablet (7 of 18), and focusing on the TV while the companion content was loading (9 of 18). Additional feedback about the companion application included increasing the time companion content was shown, which 5 participants found too short. Others suggested showing companion content during quieter parts of the programme (7 of 18).

DISCUSSION

Although we hypothesised that a longer relative time delay would result in lower immersion and affect levels, statistical analysis showed no significant differences in the immersion scores, affect scores or physiological responses. Surprisingly, even a relative time delay of 10s did not affect the participants' self-reported immersion and affect. This is mirrored in the physiological responses gathered. However, in the interviews participants reported noticing the delays and adjusting their behaviour accordingly. It is highly possible that having another screen to focus on and "sensing" the updates directly from the tablet reduced the effect of the relative time delay.

Participants were perceiving the relative time delays to be much shorter than in reality. An average estimation error of 64.2% for the 10s relative time delay condition implies that a 10s delay 'felt' only slightly more than 3s. Once we revealed the actual delay, participants were surprised that it was longer than their estimation. It is possible that the participants were so immersed in the footage that lost track of time [12], resulting in a poorer estimation score. As we collected the time estimate data from the participants retrospectively, findings using this data should be taken with caution. Our findings suggest that short time delays may not be critical when providing complementary content to factual documentaries like *Alaska: Earth's Frozen Kingdom*, as the delay between users being notified of new content and switching can effectively "overwrite" synchronisation delays of up to 3s. A more important design factor could be ensuring that the time delay does not affect the relevancy of complementary content.

We hypothesised that the longer relative time delay conditions would result in participants missing important information, leading to lower companion content recall scores. However, results show no significant effect to suggest this. There is a trend in the data to suggest lower recall scores for longer relative time delays, but it is difficult to ascertain if participants missed the screens or simply forgot what they saw. However, three participants said they were frustrated having to switch more when the companion application was loading, and five said they became impatient and lost interest when it occurred.

We aimed to investigate if participants responded to injected delays through changes in GSR and HR. However, while there was some indication that participants were negatively impacted during the 10s delay condition in the PANAS upset subscale and interviews, there were no

corresponding physiological responses. It is possible that the time delay did not bother the participants sufficiently to trigger a response even though participants commented in interviews that the delay was annoying.

When interviewed, seven participants suggested that the companion content should be presented when there was less activity on the TV. This could prevent participants getting bored by sustaining engagement, and possibly reduce the level of interruption caused by the companion content. We argue that the effects of time delay could be mitigated by careful choice of when to display companion content, and encourage content producers to be more discerning.

Limitations and future work

Better-designed companion content might cause participants to give more priority to the companion screen which might amplify the effect of the time delay. The slides used in this study were carefully designed to control for visual complexity, but participants may react differently to a more fully-featured, professionally designed companion application. Also, all participants were new to companion experiences, and may have been more enthusiastic in using both screens. This could result in increased engagement with the companion content, and higher hit rates in the recall test.

Future studies might increase exposure time of companion content to allow participants a more relaxed reading pace. Introducing interactivity, or leaving the content displayed throughout could also lead to a better understanding of how exposure time affects user experience.

CONCLUSION

This study explored how relative time delay affects companion screen experiences by analysing immersion and affect from self-report questionnaires, qualitative interview data, physiological measures, and companion content recall. Overall, no significant effect of relative time delay was found on any of these measures, though participants expressed frustration at the 10s delay in interviews. This was supported by the results of the PANAS upset subscale. A significant effect was almost detected for 10s relative delay conditions when analysing negative affect score, and some participants adopted strategies to mitigate the negative effects they experienced. Therefore, we argue that relative time delays of up to 10s should be taken seriously as they could result in a negative experience for the viewer.

Results show that the 3s time delay condition was not significantly different from the baseline condition, suggesting that such short time delays might not be critical for companion if content relevancy is intact, and may not even be perceived at all. Data from the companion content recall test indicated that participants were actively engaging with the companion content, and found no significant difference between the three delay levels. Having the main TV to focus on whenever a relative time delay occurred may have mitigated this, making the delay. Other possible

explanations include the lag in attention switching between screens and lower perceived time delay estimations.

Ultimately, users can still enjoy watching video content without companion applications. It is therefore critical that companion experiences achieve good device-TV synchronisation to ensure they improve the experience for the user, rather than becoming a source of frustration.

REFERENCES

1. Santosh Basapur, Gunnar Harboe, Hiren Mandalia, Ashley Novak, Van Vuong, and Crysta Metcalf. 2011. Field trial of a dual device user experience for iTV. *Proc.* EuroITV '11. ACM. 127-136. http://dx.doi.org/10.1145/2000119.2000145.

2. BBC R&D. Companion Screens: Creating a viewing experience across more than one screen. Retrieved January 24, 2017 from http://www.bbc.co.uk/rd/projects/companion-screens.

3. Andy Brown, Michael Evans, Caroline Jay, Maxine Glancy, Rhianne Jones, and Simon Harper. 2014. HCI over multiple screens. *Proc. CHI EA '14*. ACM, 665-674. http://dx.doi.org/10.1145/2559206.2578869.

4. Duncan P. Brumby, Helena Du Toit, Harry J. Griffin, Ana Tajadura-Jiménez, and Anna L. Cox. 2014. Working with the television on: an investigation into media multitasking. *Proc. CHI EA '14*. ACM, 1807-1812. http://dx.doi.org/10.1145/2559206.2581210

5. Moira Burke, Anthony Hornof, Erik Nilsen, and Nicholas Gorman. 2005. High-cost banner blindness: Ads increase perceived workload, hinder visual search, and are forgotten. *ACM ToCHI,* 12, (2005), 423-445. https://doi.org/10.1145/1121112.1121116.

6. David Coyle, James Moore, Per Ola Kristensson, Paul Fletcher, and Alan Blackwell. 2012. I did that! Measuring users' experience of agency in their own actions. *Proc. CHI '12*. ACM, 2025-2034. http://dx.doi.org/10.1145/2207676.2208350.

7. Empatica. Real-time physiological signals – E4 EDA/GSR sensor. Retrieved January 3, 2017 from https://www.empatica.com/e4-wristband.

8. David Geerts, Rinze Leenheer, Dirk De Grooff, Joost Negenman, and Susanne Heijstraten. 2014. In front of and behind the second screen: viewer and producer perspectives on a companion app. *Proc. TVX '14*. ACM, 95-102. http://dx.doi.org/10.1145/2602299.2602312.

9. Melanie C. Green and Timothy C. Brock. 2000. The role of transportation in the persuasiveness of public narratives. *J. Pers. Soc. Psychol.*, 79, 5, 701–721. http://dx.doi.org/10.1037/0022-3514.79.5.701.

10. Matt Hammond. A year of building upon companion screen standards. Retrieved January 24, 2017 from http://www.bbc.co.uk/rd/blog/2016-07-companion-second-screen-standards.

11. Michael E. Holmes, Sheree Josephson, and Ryan E. Carney. 2012. Visual attention to television programs with a second-screen application. *Proc. ETRA '12*, ACM, 397-400. http://dx.doi.org/10.1145/2168556.2168646.

12. Charlene Jennett, Anna L. Cox, Paul Cairns, Samira Dhoparee, Andrew Epps, Tim Tijs, and Alison Walton. 2008. Measuring and defining the experience of immersion in games. *Int. J. Hum.-Comput. Stud. 66*, (2008), 641-661. http://dx.doi.org/10.1016/j.ijhcs.2008.04.004.

13. J. Peter Kincaid, Robert P. Fishburne, Richard L. Rogers, and Brad S. Chissom. 1975. Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel. Technical Report. Naval Technical Training Command. Retrieved January 4, 2017 from http://www.dtic.mil/dtic/tr/fulltext/u2/a006655.pdf.

14. Peter J. Lang, Mark K. C. Reenwald, Margaret M. Bradley, and Alfons O. Hamm. 1993. Looking at pictures: Affective, facial, visceral and behavioral reactions. *Psychophysiology*, 30, 3 (May 1993), 261–273. http://dx.doi.org/10.1111/j.1469-8986.1993.tb03352.x.

15. Brad A. Myers. 1985. The importance of percent-done progress indicators for computer-human interfaces. *Proc CHI '85*. ACM, 11-17. http://dx.doi.org/10.1145/317456.317459.

16. Philip M. Napoli. 2012. Audience evolution and the future of audience research. *Int. J. Media Manag.*, 14, 2, 79–97. http://dx.doi.org/10.1080/14241277.2012.675753.

17. Timothy Neate, Matt Jones, and Michael Evans. 2015. Mediating Attention for Second Screen Companion Content. *Proc. CHI '15*. ACM, 3103-3106. http://dx.doi.org/10.1145/2702123.2702278.

18. Nielsen. Brain activity predicts social tv engagement. 2015. Retrieved January 23, 2017 from http://www.nielsen.com/content/dam/nielsenglobal/co/docs/Reports/2015/Nielsen%20Neuro%20Report%20April%202015.pdf.

19. Jakob Nielsen. Response times: The 3 important limits. 2014. Retrieved June 17, 2016 from https://www.nngroup.com/articles/response-times-3-important-limits/.

20. Ofcom. The communications market report. 2016. (2016)

21. Keith Rayner, Timothy J. Slattery, and Nathalie N. Bélanger. 2010. Eye movements, the perceptual span, and reading speed. *Psychonomic Bulletin & Review*, 17, 6 (December 2010), 834–839. http://doi.org/10.3758/PBR.17.6.834.

22. Jacob M. Rigby, Duncan P. Brumby, Anna L. Cox, and Sandy J. J. Gould. 2016. Watching movies on netflix: investigating the effect of screen size on viewer immersion. *Proc MobileHCI '16. ACM, 714-721. https://doi.org/10.1145/2957265.2961843.

23. Jocelyn Scheirer, Raul Fernandez, Jonathan Klein, and Rosalind W. Picard. 2002. Frustrating the user on purpose: a step toward building an affective computer Interacting with Computing, 14, 2 (February 2002), 93-118. http://dx.doi.org/10.1016/S0953-5438(01)00059-5.

24. Vinoba Vinayagamoorthy, Penelope Allen, Matt Hammond, and Michael Evans. 2012. Researching the user experience for connected TV: a case study. *Proc. CHI EA '12*. ACM, 589-604. http://dx.doi.org/10.1145/2212776.2212832.

25. Vinoba Vinayagamoorthy, Rajiv Ramdhany, and Matt Hammond. 2016. Enabling Frame-Accurate Synchronised Companion Screen Experiences. *Proc. TVX '16*. ACM, 83-92. http://dx.doi.org/10.1145/2932206.2932214.

26. Chen Wang, Erik N. Geelhoed, Phil P. Stenton, and Pablo Cesar. 2014. Sensing a live audience. *Proc. CHI '14*. ACM, 1909-1912. http://dx.doi.org/10.1145/2556288.2557154

27. Chen Wang and Pablo Cesar. 2015. Measuring Audience Responses of Video Advertisements using Physiological Sensors. *Proc. ImmersiveME '15*. ACM, 37-40. http://dx.doi.org/10.1145/2814347.2814352.

28. Chen Wang and Pablo Cesar. 2015. Physiological measurement on students' engagement in a distributed learning environment. *Psychol. Comput. Syst.*, 149–156. http://oai.cwi.nl/oai/asset/23206/23206A.pdf.

29. David Watson, Lee A. Clark, and Auke Tellegen. 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. *J. Pers. Soc. Psychol.*, 54, 6 (June 1988), 1063–1070. http://dx.doi.org/10.1037/0022-3514.54.6.1063.

30. Gillian M. Wilson and Angela Sasse. 2000. Do users always know what's good for them? Utilising physiological responses to assess media quality. In *HCI 2000*. 327–339. Springer London. http://dx.doi.org/10.1007/978-1-4471-0515-2_22.

31. Gillian M. Wilson. 2001. Psychophysiological indicators of the impact of media quality on users. *Proc. CHI EA '01*. ACM, 95-96. http://dx.doi.org/10.1145/634067.634125.