Companion apps for information-rich television programmes: representation and interaction

John Dowell
Department of Computer Science,
University College London.
j.dowell@cs.ucl.ac.uk

Sylvain Malacria
Inria, Lille, France.
sylvain.malacria@inria.fr

Hana Kim
Department of Computer Science,
University College London.
hana.kim.13@ucl.ac.uk

Edward Anstead
Department of Computer Science,
University College London.
e.anstead@ucl.ac.uk

The use of a companion app to augment viewing of information-rich television programmes is investigated. The app displays a synchronised graphical abstraction of a programme's content in the form of a concept map. Two experiments were conducted involving participants watching an astronomy documentary with the app. The first compared watching the programme with and without the app, the second compared non-interactive and interactive versions of the app. Understanding of the programme, cross-device gaze behaviour and user experience of the app were assessed. Our results show that the companion app improved participants' understanding and recall of the programme. Participants were found to manage their visual attention systematically when using the companion app and correlations were found in the way they shifted their gaze from TV screen to tablet and back in response to changes in the programme content. Increasing interaction with the app disrupted understanding of the television programme and visual attention. Participants were positive about the value of companion apps for understanding and recall of programmes but distraction and 'knowing where to look' were significant concerns.

1. Introduction

Television in the future will still be watched on large visual displays but accompanied by programme-specific apps on mobile devices. Viewers will interact with synchronised content with these companion apps, making TV a more active and participatory experience (Van Dijk, and Vos, 2001; Moulding, 2012). Companion apps are in part a response to the trend amongst viewers for using web-enabled portables while watching TV, often to perform un-related online activities (Brown *et al*, 2014). Companion apps also take advantage of the refined interfaces of smartphones and tablets, contrasting with the laborious menu navigation and generic remote control of the unpopular interactive TV.

Companion apps have the potential to support a rich variety of viewer activities including accessing additional information such as a programme's sources, creating a personalised extract of the programme, allowing users to share opinions about a programme through social media, browsing reviews, or participating in quizzes and games (Klein *et al*, 2014; Nandakumar and Murray, 2014). They may well take over from the traditional remote control for controlling the television, allowing users to further personalise their viewing of a programme. Companion apps have already appeared for a small number of TV programmes including entertainment and sporting programmes (Bruce, 2014). Possibly the first substantial evaluation of a companion app involved viewers of the 'Autumnwatch' live natural history programme using a prototype companion to access synchronised, supplementary content and provide responses to events in the programme (Jones, 2011). That evaluation concluded that companion apps can increase viewers' enjoyment of programmes yet the design factors that determine their value remain largely unknown, as do their developmental pathways.

Two primary issues confronting researchers and developers of companion apps are first, the kind of content provided by companion apps and the form in which it is presented, and second, the way users should interact with the app. In the study we report, both issues are investigated in relation to information-rich television programmes such as science documentaries. The study addresses three related questions: whether understanding and recall of a science documentary can be improved by a companion app; whether this improvement is affected by interaction with the companion app, and; whether viewers are able to manage their visual attention effectively with the additional display of a companion app. These questions were investigated through two

experiments in which participants watched a programme about astronomy whilst using a companion app to display concept maps drawn from the programme. In the first experiment participants passively viewed the companion app whilst watching the programme whilst in the second they were able to actively interact with the app. The value of the app to participants was assessed through their recall and understanding of the programme, their cross-device gaze behaviour and their self-reported experience of the app. Before presenting the study and its findings, we survey the prior related work.

2. Related Work

The convergence of personal computing technologies and television has given rise to a wholly new kind of human computer interaction (Courtois & D'Heer, 2012). Researchers have variously attempted to rationalise these new multiscreen viewing behaviours, revealing how they empower the user to "control, transfer, enrich and share" (Cesar et al., 2008) television content. Much existing work in this area has analysed the ad-hoc configuration of devices by users to fuse social networking with viewing (for example, Lochrie & Coulton 2012; Doughty *et al*. 2012). By contrast, we focus on second screen companion applications providing designed user experiences that augment television programming with additional synchronised content.

Early research recognised the potential of the television set as a high quality output device for display of graphical imagery coordinated with interactive personal digital assistants (Robertson *et al*, 1996). Recent work has built on this perspective through enhancement of broadcast television programming, distributing related content to a companion app accessed through a tablet or smartphone. These configurations offer new interactive possibilities and alternative representations of TV content on a mobile computing device. Our work focuses on the development of a companion application for information rich programmes. Such programmes have complex, sometimes abstract, content that is distilled into enlightening and enjoyable television viewing experiences. Viewers may vary greatly in their prior knowledge and understanding of the topics covered. Examples of information rich genres include science, natural history, historical drama and current affairs programming. Our work explores a companion application intended to enhance viewers' comprehension and understanding of the scientific ideas presented in the programme and to create a more engaging experience.

2.1. Content Representation on the second screen

The distribution of auxiliary content to a companion device opens a broad and unmapped range of possibilities for representing that content. Fallahkhair *et al* (2005) proposed a mobile companion application to augment language learning programmes by providing additional cultural context, translation and definition of technical terms. A prototype companion app for an episode of the Frozen Planet natural history series used combinations of images and text to provide a short summary about each animal as it appeared in the programme and an accumulating index of thumbnail images (BBC 2012). Users reported that the app: broadened and deepened their engagement with the programme; allowed a personalised viewing according to individual interests; encouraged curiosity and shared discovery; provided additional information; provided a personalised log of items of interest and items to access later; was particularly engaging for young adult viewers (who sought more information) and children (who used it to continue learning journeys); could compete negatively for their attention and was ignored during highly visual parts of the programme. They also reported that simply repeating the broadcast content without additional content was not valued and that the second screen interaction should be user-paced rather than system-paced. Nandakumar & Murray (2014) explored interactive maps to represent character relationships and plot for a crime fiction series. First-time viewers joining an episode part way through the series were found to better understand the story and the character relationships.

The question concerning the kind and form of content a companion app should provide is focused by our interest in information rich television and specifically science content, where viewer comprehension of the topics is key to determining the success of the companion. In creating a companion app for our study we assumed that it needed to provide a synopsis of the TV programme in terms of the main scientific ideas and their relationships and our companion app uses concept maps to provide this synopsis. Concept maps have long been used as a teaching and learning technique in science education (Novak, 1990) and three decades of research into concept maps as an educational tool show they benefit learning across a wide variety of settings used in different ways (Nesbit and Adesope, 2006). In particular, animated concept maps that visually grow have been shown to benefit learning over static concept maps (Blankenship & Dansereau 2000). It has also been shown that concept maps enhance learning when used in combination with a spoken recording (Adesope & Nesbit,

2013); it was an open question whether this benefit would also occur with a television programme competing for the viewer's visual attention.

2.2. Interaction with the second screen

Companion applications offer users diverse opportunities to interact with television programming on secondary devices without obscuring the big screen. Examples include companion app programme guides (Cruickshank *et al*, 2007 and Park *et al*, 2006) and bespoke experiences designed to enrich viewing of individual programmes or general viewing. Supplementary content created and shared by viewers themselves has also been explored; the FanFeeds prototype companion app enabled viewers to search for and attach videos and web documents relevant to a particular programme (Basapur et al 2012). This work built on the earlier development of a second screen application that enhanced programming with auxiliary synchronised information updates from third party sources such as wikipedia and news outlets (Basapur et al., 2011).

How users are able to manage their visual attention over both screens is recognized to be a key issue for the uptake of companion applications (Geerts *et al*, 2014). High cognitive demands on users interacting with multiscreen television are assumed to diminish understanding of the programme's content and the enjoyment of watching (Vatavu & Mancaş, 2014). Yet prior investigation of this issue is limited. Holmes et al (2012) reported that 30% of the viewer's total gaze is directed at the companion app and that interactive content draws still more attention. Brown et al (2014) found that only 20% of total gaze was directed at a companion app.

Previous work has also examined visual attention simply to television programmes when no second screen is present. The duration of 'looks' at programmes was examined with different lengths of looks interpreted as reflecting processes of attention and comprehension (Hawkins *et al*, 1997). Four distinct durations have been characterised as: monitoring looks (less than 1.5 seconds); orienting looks (1.5 to 5 seconds); engaged looks (6 to 15 seconds); and stares (longer than 15 seconds) (Hawkins *et al*, 2005). Factors that contribute to the initiation, maintenance and termination of looks have been examined (Burns and Anderson, 1993). An attentional inertia in viewers' gaze shifts has been found such that viewers exhibit a tendency to continue rather than shift their gaze (Anderson *et al,* 1979).

Cues that affect gaze shifts in television viewing have also been investigated. Even young children are seen to respond systematically to cues about the comprehensibility and informativeness of television content (Anderson *et al,* 1979), suggesting that television viewing elicits active processing of content and is far from being a passive activity (Collins, 1982). Moreover, viewers rapidly decide whether to continue watching on the basis of salient cues with learned associations with content (Huston and Wright, 1983). Content features have been investigated to understand how they may cue changes in visual attention, features including female characters, male voices, auditory changes and sound effects, activity or movement, camera cuts, zooms and pans, laughing and applause, and still pictures (Collins, 1982). In the future, both television programmes and companion apps may be designed with implicit cues and content transitions that give users opportunities to shift their gaze at will. Alternatively, viewers might be given explicit notifications by the programme and the app about when to shift their gaze (Neate, Jones et al. 2015).

Interactions with a companion app are examined in our study. The use of the app as a passive display of secondary information is compared with its active use when interactive information and graphics re-sizing functions are available. The management of visual attention across both screens is examined in relation to the duration of gaze, gaze shifts and cues influencing those shifts. The experience of watching two screens together is also investigated.

3. Method

The method and design of both experiments are described including the development of our prototype app we named 'CompanionMap'.

3.1. Programme and concept maps

Programme

A documentary about astronomy ('The seven ages of starlight', BBC) was selected for use in the studies because of its substantial conceptual and factual content. The 90 -minute programme had been made for a broad audience and screened (on BBC4) some two years prior to our study. Rather than using the complete programme, two seven minute recorded 'episodes' were chosen for the studies on the advice of one of the astrophysicists who had worked with the original production team. The episodes were chosen for their balance of clarity and complexity, one episode concerned supernovae, the other black-holes.
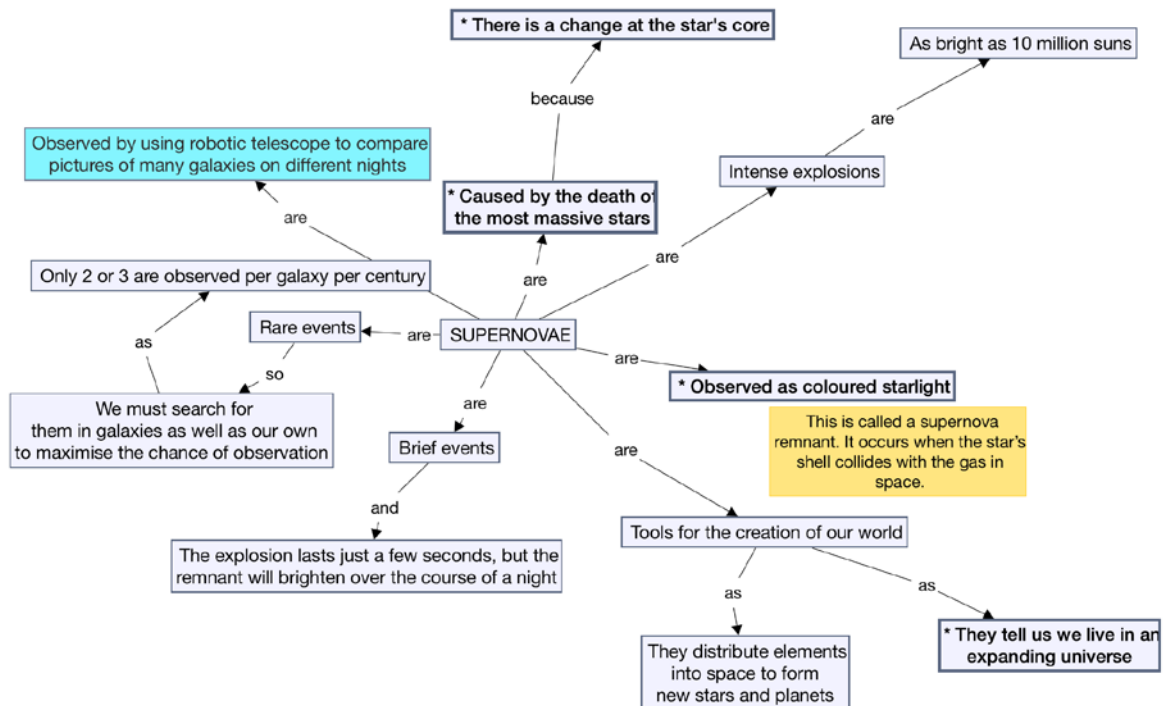
Concept maps

A concept map was created for each of the selected episodes by a graduate physicist using the Cmap editor (Canas et al, 2004). The maps contained primary concepts explicitly presented in the episodes and their relationships. Concepts were expressed in the map by single terms, propositions or short sentences (e.g. "intense explosions", "as bright as 10 billion suns", etc). Each map consists of some 20 concept boxes judged by the physicist to be critical to the programme episode; the maps contain most but not all of the primary concepts in the episode. The kinds of relationships between the concepts were reduced to a minimal set (e.g., "are", "and", "which is", etc) in order to make the maps most readable. The nodes and links conformed with the syntactic convention that allows the construction of new propositions from two nodes and the link that joins them (e.g., /intense explosions /are /as bright as 10 billion suns, etc). The supernovae map created is shown in Figure 1.

3.2. The companion app

CompanionMap is the app we created to display concept maps synchronised with programmes. The app reads concept map XML files created with the Cmap editor and displays them in coordination with the programme, progressively revealing their content. New concepts appear gradually fading-in and with a different colour to aid visual search (avoiding the "map shock" (Adesope & Nesbit 2013) experienced when trying to read a complete concept map all at once). The node loses its colour once the relevant passage of the programme episode has passed and another node fades-in with a colour highlight. Hence, the maps grow continuously until the end of the episode. The map is eventually visible on the screen in its entirety and neither panning nor zooming are provided. The text on the map is in an 8-point sans serif font. Figure 1 shows the supernovae concept map displayed on CompanionMap.

In its base mode of use, CompanionMap is simply a dynamic display without interaction. An interactive variant was created for Experiment 2 to enable viewers to access additional information about particular concepts. Concepts with additional information are visually marked with bold text and an asterisk. By tapping on these nodes, additional text appears in a dedicated 'tooltip' note providing further information and explanations. Interactive notes (of which there were 10) can be dragged around if they occlude relevant parts of the map and dismissed by tapping on the map canvas. A second kind of interaction is the shrinking of concept nodes using an extended press gesture.

CompanionMap is implemented as an iOS app built in CocoaTouch ObjectiveC and run on an iPad. The additional content for the interactive notes is loaded dynamically from a dedicated text file. The associations to the relevant concepts was hardcoded in the XML description of the concept map. CompanionMap communicates via sockets with a dedicated video player that continuously sends the position of its scrubber in the TV programme so that CompanionMap can update its display.

Figure 1. Concept map created from the supernovae episode of The Seven Ages of Starlight and displayed on CompanionMap. The node with fill colour is the one most recently to appear. A yellow interactive note is also shown.

3.3. Measures and Scoring

Probe questions: measures

A set of twenty questions was created for each episode to probe viewers' prior knowledge and their understanding and recall of the programme. Since the questions would be used to assess the impact of the concept map, questions of two kinds were created. Some questions probed knowledge of part of the TV programme episode where the central concepts also appeared on the concept map (Companion-Programme-Questions, or 'CPQ's). Other questions probed knowledge of some part of the episode and related content but where no specific concept was added to the map (Programme-Questions, or 'PQ's). Nine PQs and eleven CPQs were created for each episode. The questions were all of multi-choice type and six examples appear in Figure 2 with the answers underlined. The questions were created by the graduate physicist who had created the concept maps and were modified through pilot testing to ensure a moderate level of difficulty.

In Experiment 2, participants used the interactive version of CompanionMap which includes 10 interactive notes providing additional information about stars not contained in the documentary. To measure participants' understanding of this additional information, an additional eight questions probing knowledge about the interactive notes was added.

Probe questions: scoring

In both experiments, participants' knowledge of the content of the TV programme episodes was assessed using the probe questions. Participants' were asked to answer the same probe questions before as well as after watching each episode, to confirm that there was substantial learning from the programme. The two question types provided two dependent measures: the number of correct/incorrect/null/ answers to the CPQs would measure knowledge of those parts of the TV programme that had a corresponding concept on CompanionMap; the number of correct/incorrect/null/ answers to the PQs would measure knowledge of those parts of the TV programme that had no corresponding concept on CompanionMap. In Experiment 1, participants would use CompanionMap in only one of two trials but nevertheless would answer the CPQs for both trials. Hence, these dependent measures would expose the effect of using CompanionMap on knowledge about the programme displayed on the app.

---

1. When does a supernova occur?
   - When a new star is created.
   - When one of the oldest stars dies.
   - <u>When one of the most massive stars dies.</u>
   - I don't know.

2. Why does a supernova occur?
   - The star explodes because it experiences a runaway nuclear reaction inside.
   - <u>Its core collapses because it runs out of fuel.</u>
   - The star collides with another object in space.
   - I don't know.

3. A supernovae is as bright as:
   - Our sun.
   - 1 million suns.
   - <u>10 billion suns.</u>
   - I don't know.

4. Someday our sun will explode as a supernova.
   - True.
   - <u>False.</u>
   - I don't know.

5. What can you see left behind after a supernova?
   - A cloud of yellow dots in the sky.
   - <u>A residue of many colours.</u>
   - We cannot see the after effects of a supernova.
   - I don't know.

6. Different elements in the supernova remnant appear as:
   - Different sizes.
   - Different shapes.
   - <u>Different colors.</u>
   - I don't know

Figure 2. Sample of the twenty probe questions created to measure understanding and recall of supernovae episode of The Seven Ages of Starlight.

Free-recall: measures

As further investigation of their learning, participants were asked to state verbally all the ideas from the programme that they could freely recall. These statements would be recorded and later transcribed. They would then be parsed manually and categorized to identify the distinct concepts recalled by each participant. Those concepts would be further categorized as ones referenced on CompanionMap.

Free-recall: scoring

The number of concepts participants could recall and their accuracy were assessed. The method for marking the recall was adapted from (Nesbit & Adesope 2011). A listing of the primary concepts in each programme episode (45 in black-holes, 38 in supernovae) was used to score participants' recall. Partially correct or partially complete statements were scored as 1, correct and complete statements were scored as 2.

User satisfaction: measures and scoring

Participants' satisfaction was assessed by qualitative and quantitative measures using a usability test questionnaire and semi-structured interview, respectively. The usability test questionnaire measured 'efficiency', 'satisfaction', 'perceived usefulness' and 'learning object evaluation' using the 5-point Likert scale were created. In semi-structured interviews, opinions about the advantages and disadvantages of CompanionMap were elicited and suggestions for modifying the app's design.

Gaze transitions: measures

Measures of eye gaze were obtained from the video-recordings of participants' faces. Gaze direction towards the TV screen, tablet or off-screen were recorded in terms of the time at which changes of gaze direction occurred and the direction of the change. The main aim was to discover how participants' gaze transitions between screens were affected by content changes in the documentary. A total of 32 participants' gaze data was aggregated from the trials with the base version of CompanionMap trials in both experiments.

Gaze transitions: scoring

Eye gaze was investigated in relation to changes in both auditory content and visual content of the episode. It was assumed that these content changes would be the primary determinants of gaze shifts opposed by the attentional inertia earlier discussed (Anderson et al, 1979). Auditory content changes would be used to determine when participants move their gaze from companion app to TV screen. Visual content changes would be used to determine when participants move their gaze from the TV to the companion app on the assumption that that participants would try to make the best use of their visual attention resources by looking away from the television screen when it was not providing useful or interesting visual information. We therefore assumed that auditory content changes would not directly determine gaze shifts from television to companion

Each episode video was analysed to determine at which points the auditory content changed between being either non-verbal sound, a narrator's voice, or a definite person speaking (e.g., an astrophysicist talking about their work). So auditory content changes had six types: from non-verbal to narration, from non-verbal to interviewee, from narration to interviewee, from narration to non-verbal, from interviewee to narration, and from interviewee to non-verbal.

Each episode was also analysed to determine at which points the visual content changed between being either filler (i.e., non-specific and uninformative imagery acting as a visual link), illustration (a specific visual image of some spoken verbal content), or portrait (an image of a person, usually an interviewee since the narrator was never visible). Thus, visual content change had six types: from filler to illustration, from filler to portrait, from illustration to portrait, from illustration to filler, from portrait to illustration, and from portrait to filler.

We measured three gaze shifts driven by the auditory content changes: 1) from CompanionMap to TV 2) from CompanionMap to CompanionMap (i.e., no shift) 3) from CompanionMap to Off-screen. We measured three gaze shifts driven by the visual content changes: 1) from TV to TV (i.e., no shift) 2) from TV to CompanionMap 3) from TV to Off-screen.

Gaze shift indices were computed as proportions of the number of each kind of origin-to-destination transition in relation to the total number of transitions from that origin; for example the total number of TV to CompanionMap shifts divided by the total number of gaze shifts from the TV (i.e., the sum of TV to TV shifts, TV to CompanionMap shifts and TV to Offscreen shifts).

Eye gaze-frequency and duration: measures and scoring (second experiment only)

Experiment 2 included further comparisons of gaze frequency and duration with the base line and interactive CompanionMap versions, since both trials involved the companion app unlike experiment 1. Gaze frequency and duration towards the TV screen, tablet, or off-screen were measured from the video-recordings of the participants' faces.

3.4. Participants

All participants were recruited from within the university and were students and members of the research staff. Participants were screened to ensure they were fluent in English, that they had not seen the Science TV programme *Seven ages of Starlight* and that they did not have advanced knowledge of astronomy. They were required to be people who did not need spectacles as this would obscure their direction of gaze in video-recordings we would make during the experiment. For Experiment 1, 16 participants were recruited, 9 males and 7 females, all ranging in age from 23 to 62, with a mean age of 32.5 years (SD=10.7). Participants were allocated randomly to conditions.

For Experiment 2, 16 participants were recruited and included 9 males and 7 females. The participants ranged in age from 19 to 41 with a mean age of 30.8 years (SD=5.98). Participants were allocated to conditions randomly.

3.5. Apparatus

Both experiments were conducted in a testing room with subdued lighting and containing a 40" high definition television mounted on a stand, an armchair set in front of the TV and a low table adjacent to the chair. The recorded programme episodes were played from a laptop (MacBook pro) by HDMI connection to the TV and using a dedicated video-player. To their left on the low table, a second laptop presented participants with the set of probe questions to answer after watching the programme episode. The questions were presented via the Qualtrics survey Platform. In both experiments a video camera was used to record the participant's face in order to determine eye gaze, free-recall and interview.

3.6. Design

Experiment 1 used a one-way repeated measures design comparing Condition (watching the TV programme with and without CompanionMap). Each participant viewed both programme episodes, one with CompanionMap; the order of the episodes and the use of CompanionMap was counter-balanced in a latin square of four groups to avoid learning effects. The probe questions were administered both before and after watching the programme episodes. Experiment 2 also used a one-way repeated measures design comparing the use of the interactive version of CompanionMap with the base version. The experimental design was identical to the first experiment except for Condition (non-interactive and interactive CompanionMap). All dependent variables were identical.

Figure 3. Participants viewed CompanionMap on a tablet they were holding while watching the television programme on a large screen. The laptop to one side displayed the probe questions after each episode and logged participants' responses.

3.7. Procedure

Although the procedures differed slightly in the different trials in the two experiments, the overall process was the same except for the order of treatments. The procedure for the first condition began with participants completing a questionnaire about themselves then answering the probe questions for the episode assigned to that trial. To encourage them to answer accurately, participants were told a scoring system was in place that would penalise incorrect answers, give a zero for null answers but reward the highest participant score with a bonus payment. After watching the programme episode, the participant was asked to verbally recall as many ideas or facts as they could from the video recording. Participants then answered the probe questions once more, followed by the TV viewing questionnaire. The experiment then continued with the second trial using the same procedure as the first. In the trial with CompanionMap, the participant was introduced to it with an unrelated concept map. They were encouraged to try to monitor CompanionMap continuously while watching the programme episode. The participant watched the video and performed the recall test, answered the probe questions once more and then the questionnaire. The trial completed with a short semi-structured interview. In Experiment 2, the same procedure was followed except that participants were able to explore the manipulation of a sample concept map with CompanionMap before watching the programme episodes.

4. Results

4.1. Experiment 1

Programme understanding and recall

The mean number of correct answers to the probe questions before watching the programme episodes was 6.4 rising to 13.7 after the programmes were viewed. The number of correct answers to the CPQs after watching the programmes was greater (T=2.2, p=.044) when CompanionMap had been used (9.06, sd=1.43) than when it had not (8.375, sd=2.02). Further, Cohen's effect size value (d = 1.12) suggested a moderate to high practical significance.

No significant differences were found in the numbers of incorrect (T = -0.46, p=0.65) and null answers (T= -2.06, p = 0.057) to the CPQs in the two conditions. No significant difference (T (15) =0.112, p= 0.912) was found between conditions in the number of correct answers to the PQs after watching the programme. No significant differences were found in the time taken to answer the CPQs and the PQs (T (15) = 0.103, p=0.919). Free recall scores showed no significant differences (T (15) = 0.684, p=0.504) between CompanionMap (mean = 9.13, SD=3.90) and non-CompanionMap trials (mean = 8.31, SD=4.7) on concepts that appeared on the concept maps
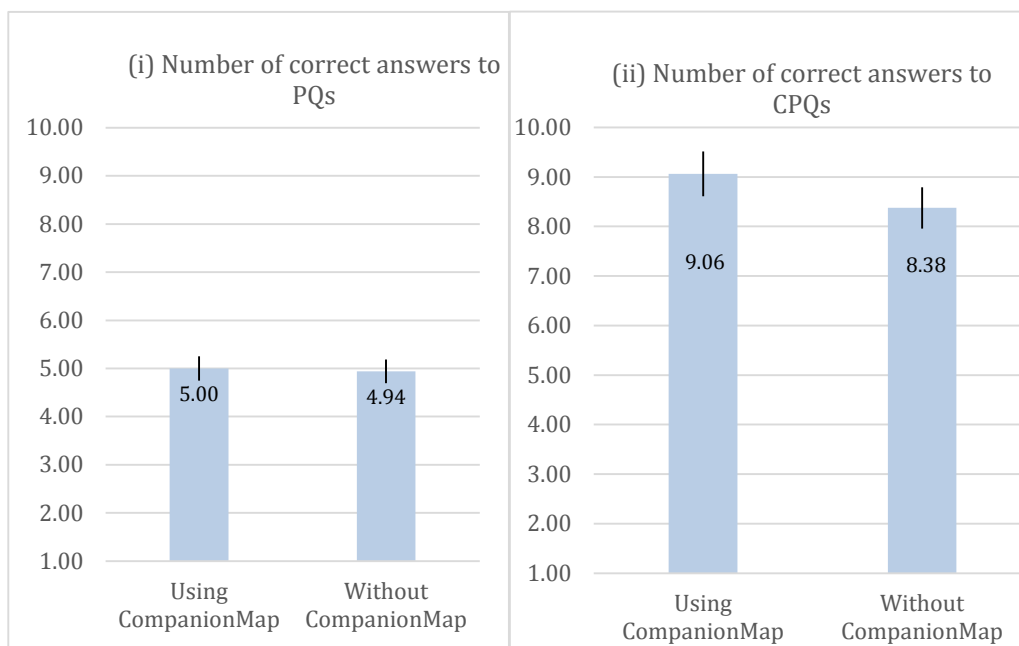
Figure 4. (i) Number of correct answers to PQs (questions on parts of the programme without corresponding concepts on the CompanionMap), with and without using the companion app. (ii) Number of correct answers to CPQs (questions where concepts were on the companion app) with and without using the companion app. (Error bars: 5%).

User satisfaction

Participants' reports after each trial showed no overall difference in satisfaction with watching the TV with CompanionMap (T (15) = 0.148, p= 0.884). They reported levels of satisfaction with CompanionMap on a five point Likert scale as: usefulness at focusing (3.32); understanding (3.88); remembering (4.0); aesthetics (3.19). Users expressed a wish for the companion app to provide more interaction (3.59).

In interviews after completing Experiment 1, 37% of participants stated strongly that they would use the concept map companion app in the future and a further 44% said they would use it depending on the situation. Participants reported that the app assisted memorization, concentration and understanding. They also reported that it was distracting and they often did not know where to look. They suggested that the app could be

improved by additional information and by controls over the television for pausing and reviewing the programme.

4.2. Experiment 2

Programme understanding and recall

The number of correct answers to the CPQs after watching the programmes was greater (T(15) =3.955, p=0.001) when the base version of CompanionMap was used (mean = 8.94, SD = 1.77) than when the interactive version was used (mean = 7.19, SD = 1.80). No difference was found between conditions in PQ scores (T (15) =-0.324, p = 0.751). The proportion of correct answers to questions about the additional information available on the interactive notes (mean = 0.52, SD = 0.22) in the interactive version was lower (T(15) = 2.65, p=0.018) than the proportion of correct answers to the CPQs (mean = 0.65, SD = 0.16).

User satisfaction

User satisfaction was greater (T (15) = -3.381, p=0.004) with the base version of CompanionMap (mean = 31.25, SD=4.33) than with the interactive version (mean = 29.06, SD=3.53). Many more participants also stated that they preferred the base version of CompanionMap (75%) to the interactive version (19%). The most commonly given advantage of the base version was that it was less distracting (60%). The most commonly given advantages of the interactive version were its additional information (67%) and interactive functions (33%); the most commonly stated disadvantage was that it was distracting (94%). Participants also stated a wish to have a sequence order displayed (e.g., a visual trail across nodes), a pause function and a highlighting function.

Interactive notes

Of the total of 10 available interactive notes, the mean number opened was 7.31 (SD =2.52) and the mean number of total opening and closing actions was 24.19 (SD= 11.12).

Gaze shifts with programme cues

Gaze shifts between companion app and TV screen in response to cues in the programme are now reported and are based on the aggregation of eye gaze data collected in Experiments 1 and 2 from the trials with the base version of CompanionMap, a total of 32 participants.

Gaze shifts with auditory content changes. Indices for gaze shifts from the CompanionMap to the TV in response to auditory content changes were: 84.9% for 'non-verbal to narration; 93.6% for 'nonverbal to interviewee'; 88.8% for 'narration to interviewee'; 100% for 'interviewee to narration'; 62.7% for 'narration to nonverbal', and; 48.3% for 'interviewee to non-verbal'. These data are shown graphically in Appendix 1.

Gaze shifts with visual content changes. Indices for gaze shifts from the TV to CompanionMap in response to visual content changes were: 73.1% for 'filler to illustration'; 73% for 'illustration to filler; 53.6% for 'filler to portrait'; 43.2% for 'illustration to portrait'; 48.8% for 'portrait to illustration', and 58% from 'portrait to filler'. These data are shown graphically in Appendix 2.

Gaze durations

The proportion of total gaze durations by screen in each condition were:

Base version of CompanionMap: 71.7% to TV; 27.8% to CompanionMap; 0.5% to off-screen.

Interactive version of CompanionMap: 47.1% to TV; 50.5% to CompanionMap; 0.4% to off-screen.

With the interactive CompanionMap, the mean duration of gaze at the TV was 7.6 seconds (SD=4.0) and at CompanionMap it was 5.3 seconds (SD=2.5). With the base version of CompanionMap, the mean duration of gaze at the TV was 5.49 seconds (SD= 1.70) and at CompanionMap it was 2.84 seconds (SD= 0.84).

5. Discussion and conclusions

This study investigated whether companion apps can support viewers' understanding and recall of science programmes, whether users are able to manage their visual attention effectively with the additional display of a companion app, and how adding interactivity to the companion app affects these issues. It also sought participants' views on the experience of using the companion app. Answers to each of these questions evident from the two experiments are now discussed.

5.1. Understanding and recall

Participants in Experiment 1 achieved higher scores with the CPQ probe questions after using CompanionMap than after simply watching the programme episodes without the companion app. CPQs probed knowledge about concepts in the programme that also appeared on the concept map. The results show that the concept map helped participants to better understand and recall the programme episodes. When the questions referred to concepts that had been in the programme but not on the concept map (PQs), participants' scores were the same whether or not they used CompanionMap. This indicates that the effect of the companion app on understanding and recall was due to the contents of the concept map and its representational form, rather than being an indirect effect of simply using the tablet. The expected marked rise in their before and after scores is evidence of participants' substantial learning about the life of stars from the programme episodes. So it was within this highly active learning process that CompanionMap achieved its impact, particularly on gaining new knowledge as evidenced by the non-significant differences in incorrect and null answers across conditions.

The effect of CompanionMap on recall and understanding would likely have been greater if the probe questions asked before and after each episode had been different, therefore removing the cueing of those parts of the programme during viewing. The effect might also have increased if participants had been asked to watch more of the programme.

The free-recall results did not achieve significance so did not provide additional evidence of the benefits of CompanionMap. The reliability of the free recall method was compromised by the requirement for participants to verbalise their recall to the experimenter rather than write their recollections alone, and by the marked variability in the quantity of recall between participants.

The advantage found in Experiment 1 for using the CompanionMap while watching the television programme extends previous findings about concept maps (Nesbit and Adesope, 2006). Animated concept maps have been previously shown to have a benefit over static maps in a setting where no other media are present (Blankenship & Dansereau 2000). Animated concept maps have also been found to enhance learning when used in combination with a spoken recording (Adesope & Nesbit, 2013). Given that watching television is visually demanding and that an additional concept map increases the demands on visual attention, the benefit could not be taken for granted and a poorly contrived map might well interfere with visual attention and learning.

The measures of understanding and recall used in this study are similar to those used in studies of educational media. There is clearly an open question about how learning, as represented by these measures, relates to the learning experienced when a person watches a science documentary for interest and pleasure. By using experimental methods in a laboratory rather than unobtrusive observation in viewers' homes, the studies clearly compromised ecological validity to gain reliability in the findings.

5.2. Effect of interactivity

The interactive version of CompanionMap provided additional information about particular concepts; tapping a concept would make an interactive note appear. Participants could also shrink nodes to keep track of ones they had viewed. Participants in Experiment 2 typically looked at most of the interactive notes and appear to have either opened and soon closed the information, or left it open while looking away. Participants achieved better scores for the CPQ probe questions with the non-interactive version of CompanionMap. We therefore conclude that interaction with the companion app was interfering with participants' understanding and recall of the television programme. No difference was found in PQ scores with and without the interactive version, indicating that the interactivity was specifically harming the advantages conferred by the concept map. Scores on CPQ probe questions were higher than scores on questions about the interactive notes, though this effect was confounded by the interactive notes adding to rather than simply reinforcing the programme content.

5.3. Managing visual attention

How viewers will manage looking at an additional screen of information while watching a television programme is the most controversial and poorly understood issue surrounding companion apps. Our experiments showed that participants made very active use of the companion app. When the base version of CompanionMap was used, participants viewed the companion app some 40% as much as the television screen; with the interactive app, their gaze was equally divided with the television screen. Adding interactivity to the

app clearly makes viewers spend more time looking at the app and less time at the television programme. These data substantially extend previous findings about the proportion of total gaze directed at a companion app.

Participants' visual attention was marked by the high number of transitions between screens (as many as 40 or 50 for some participants in a single programme episode) and the relatively short periods of their looks at those screens. Looks at the TV were twice as long as looks at CompanionMap, a proportion that declined with the interactive version when the duration of looks at both TV and tablet lengthened. These results, taken together, show that participants were finding the companion app as visually significant for them as the television programme and they were actively managing their visual attention. The experiments examined the cues that participants use to determine when to shift their gaze from one screen to the other. The analysis of gaze shifts in response to auditory content changes in the programme showed a very strong tendency to look back at the television when a voice starts speaking or when there is a change in speaker. Non-verbal audio changing to a speaker, whether it is a narrator or interviewee, will often cue the viewer to look from the companion app to the television. When the voice changes from narrator to interviewee or vice versa there is an even stronger apparent tendency to look back at the television.

Gaze shifts from television to companion app cued by changes in the visual content of the programme have a much less clear pattern. The balance of shifts in response to visual cues is evenly matched in our data. Nevertheless it was evident that participants tended to shift their gaze (73%) from the programme when a visual illustration changed to visual filler, that is, when informative content changed to non-informative content. Interestingly, they also shifted (73.1%) their gaze to the companion app when visual filler was replaced by a visual illustration in the programme. This could be interpreted as participants seeking a verbal label for the illustration when narration is not provided. This interpretation is supported by the absence of such a shift tendency (45%) when portrait becomes illustration, i.e., when the programme shows some visual material after an introduction by a person visible on screen. The uncertainties involved in our coding of the visual content data (e.g., when an illustration becomes filler) may be obscuring clearer correlations with the gaze shift data. Our assumption that gaze shifts from television to companion app are visually driven but not auditorially-driven is likely to be too simple, particularly if as has previously been claimed (Basil, 1994), the auditory channel dominates the visual channel when watching factual television programmes. Participants' clear reports of their difficulties in knowing where to look are the most likely explanation for the absence of overall patterns in gaze shift correlations with visual content changes.

5.4. User experience

Participants gave positive scores across the factors in our questionnaire in Experiment 1, particularly for the benefits of understanding and recall of the programme. However there was no significant preference for using CompanionMap or for watching the programme episodes without it in Experiment 1. The majority of participants in Experiment 1 said they would choose to use CompanionMap, other things being equal. Our interviews with those participants revealed a main concern with the distraction created by the app and the difficulty of knowing where to look. Their preference for CompanionMap to provide more information appears to be inconsistent with their comments about distraction, a further fruitful area for research. Participants in experiment 1 also said they would like the companion app to be more interactive, but participants disliked the interactive version of CompanionMap in Experiment 2 and the overwhelming reason given was the greater distraction it caused. Whether an alternative form of interactivity would be more satisfactory is unclear and requires further investigation. In interviews conducted at the end of each experiment, participants expressed a wish to see programme sequence order added to the concept maps, to have control over pausing the TV programme from the concept map and to have a tool for adding highlights to the maps.

5.5 Design implications

Viewers may often choose to watch a science documentary or similar information-rich programme because they will learn about a subject that interests them. A central purpose of companion apps in that context is to support the learning process and our study with CompanionMap indicates how that can be achieved.

The representation provided by a companion app should be a synoptic reinforcement of the television programme's content so that viewers can more easily recognise the main facts and ideas in the programme and how they are related. This representation should be the means for categorizing, indexing and therefore re-finding content already viewed. CompanionMap provides a synoptic representation in the form of concept maps

and these were found to increase learning. As well as synoptic reinforcement, companion apps should allow viewers to access additional related information, as CompanionMap does with its interactive notes. How much additional information and how it should be accessed are currently not well understood.

The representation provided by a companion app also needs to be a persistent view of the parts of the programme already seen. Television is, of course an intrinsically transient medium and viewers' recall of what they have watched needs to be augmented since information-rich programmes often involve relating together their parts within a linear narrative and drawing together the ideas. CompanionMap's concept maps remain visible continuously. The representation provided by a companion app should also be synchronised with the programme content currently viewed to help viewers see and read the right accompanying content at the right time, that content accumulating dynamically. CompanionMap uses animation to add map nodes synchronised with the programme with colouring and fading animation effects to indicate the newest nodes and nodes appearing. Synchronisation with the programme is definitional for an effective companion app and the dynamic accumulation of content makes that synchronisation visible.

The representation provided by the companion app needs to be coordinated with the programme. Coordination is the designed combination of contents of both television programme and companion app to take account of the viewer's focus of attention. Viewers will find their own ways of dividing and managing their attention between screens, for example, by looking back to the TV when a different voice is heard as we found in our experiments. By implication, visual attention switching should be self-paced rather than system-paced with viewers deciding when to look to the other screen and having controls to pause the programme and app. CompanionMap uses a relatively slow accumulation of map content and in our studies users were able to maintain their attention to the programme whilst making frequent, brief gaze shifts to the app. To fully achieve coordination requires that both the programme and companion app contain opportunities and cues for the viewer to shift their gaze, meaning that the production of programme and app needs to be combined and at the same time.

Interaction will increase the visual attention given to the companion app in terms of the length of looks and the balance of attention given to the app. As was seen in the second experiment with CompanionMap, understanding of the programme can be harmed by interaction and viewers report that greater interactivity is distracting. To prevent additional interaction adversely affecting attention to the programme, positive design compensations are needed to reinforce the programme content and direct attention back to it, for example, selectively hiding interactive content on the app to coordinate with the programme content.

5.6. Conclusions and final comment

Our study has shown that companion apps have the potential to improve viewers' understanding and recall of information-rich programmes. The concept map representation of programme content was found to improve participants' learning from the science documentary. Our study has also shown that users are able to manage their visual attention systematically and perhaps effectively when using the companion app. Interactivity is clearly a serious design issue for companion apps. A tablet or smartphone is by its nature a highly interactive device but more interaction with our companion app interfered with participants' understanding of the television programme, it further disrupted their visual attention and it was relatively disliked.

Television is changing as it converges with the internet and companion apps are likely to become a routine part of watching programmes. The emergence of companion apps has been largely driven by the technical possibilities, many of which remain only partially visible. To date, the development of companion apps has been a process of pure invention. Research needs to set out the design space for companion apps and empirically establish the design constraints, user needs, interaction methods and user interface frameworks; this will involve a combined effort of researchers in media design, media technology and human computer interaction. The study we have reported makes clear viewers' enthusiasm for companion apps and the value of those apps for enhancing viewers' understanding of and engagement with science programmes.

References

Adesope, O. O., & Nesbit, J. C. (2013). Animated and static concept maps enhance learning from spoken narration. *Learning and Instruction*, *27*, 1-10.

Anderson,D., Alwitt, L., Lorch, E., Levin, S. (1979). Watching Children Watch Television. *Attention and Cognitive Development*, 331-361

Basapur, S., Harboe, G., Mandalia, H., Novak, A., Vuong, V., & Metcalf, C. (2011). Field trial of a dual device user experience for iTV. *Proceesings of EuroITV '11 ACM*.

Basapur, S., Mandalia, H., Chaysinh, S., Lee, Y., Venkitaraman, N., & Metcalf, C. (2012). FANFEEDS: evaluation of socially generated information feed on second screen as a TV show companion. *Proceedings of EuroiTV '12 ACM*.

Basil, M. (1994). Multiple resource theory 1: Application to television viewing. *Communication research*, 21(2), 177-207

BBC (2012) The frozen planet dual screen experience. http://companion.prototype0.net/examples/frozen-planet-companion-2012. Accessed June 17, 2015.

Belin, P., Zatorre, R., Lafaille, P., Ahad, P., and Pike, B. (2000) Voice-selective areas in human auditory cortex, *Nature 403*, 309–312.

Blankenship, J., & Dansereau, D. F. (2000). The effect of animated node-link displays on information recall. *The Journal of Experimental Education*, *68*(4), 293-308.

Brooke, J. (1996). SUS- A quick and dirty usability scale. In Jordan, P., Thomas, B., Weerdmeester, B., and McClelland (Eds), *Usability evaluation in industry* (pp.189-). London, UK: Taylor & Francis

Brown, A., Glancy, M., Evans, M., Jones, R., Jay, C., Harper, S. (2014). HCI over multiple screens. *CHI '14 Extended abstracts on human factors in computing systems*, 665-674

Bruce, F. (2014, March, 18) Antiques Roadshow to launch BBC's first ever smartphone companion app, http://www.bbc.co.uk/mediacentre/latestnews/2013/antiques-roadshow-app.html Accessed June 17, 2015.

Burns, J., and Anderson, D. (1993). Attentional inertia and recognition memory in adult television memory. *Communication research*, 20(6), 777-790

Cañas, A., Hill, G., Carff, R., Suri, N., Lott, J., Eskridge, T., Gomez, G, Arroyo, M. and Carvajal, R. (2004). CmapTools: A knowledge modeling and sharing environment. *Concept Maps: Theory, Methodology, Technology Proc. of the First Int. Conference on Concept Mapping*, 1, 125-133

Cesar, P., Bulterman, D. C. A., & Jansen, A. J. (2008). Usages of the secondary screen in an interactive television environment: control, enrich, share, and transfer television content. *Changing Television Environments*, 5066, 168–177.

Collins, W. (1982). Cognitive processing in television viewing. In Pearl, P., Bouthilet, L., and Lazar. J (Eds), *Television and behaviour: Ten years of scientific progress and implications for the eighties (pp.9-23),* U.S.A: National Institute of Mental Health)

Courtois, C., & D'heer, E. (2012). Second screen applications and tablet users: constellation, awareness, experience, and interest. *Proceedings of EuroiTV '12 ACM*.

Cruickshank, L., Tsekleves, E., Whitham, R., Hill, A., & Kondo, K. (2007). Making interactive TV easier to use: Interface design for a second screen approach. *The Design Journal*, 10(3), 41–53.

Doughty, M., Rowland, D., & Lawson, S. (2012). Who is on your sofa?: TV audience communities and second screening social networks. *Proceedings of EuroiTV '12* ACM.

Fallahkhair, S., Pemberton, L., & Griffiths, R. (2005). Dual Device User Interface Design for Ubiquitous Language Learning: Mobile Phone and Interactive Television (iTV). In *Proceedings of the IEEE International Workshop on Wireless and Mobile Technologies in Education* IEEE.

Geerts, D., Leenheer, R., De Grooff, D., Negenman, J., & Heijstraten, S. (2014). In front of and behind the second screen: viewer and producer perspectives on a companion app. *Proceedings of TVX'14 ACM*.

Hawkins, R., Pingree, S., Bruce, L., and Tapper, J. (1997). Strategy and style in attention to television. *Journal of broadcasting and electronic media*, 4, 245-264

Hawkins, R., Pingree, S., Hitchon, J., Radler, B., Gorham, B., Kahlor, L., Gilligan, E., Serlin, R. Schmidt, T., Kannaovakun, P., and Kolbeins, G. (2005). What produces television attention and attention style? Genre, situation, and individual differences as predictors. *Human communication research*, 31 (1), 162–187

Holmes, M., Josephosn, S., and Carney, R. (2012). Visual attention to television programs with a second-screen application. *Proceedings of the symposium on eye tracking research and applications*. 397-400

Huston, A., and Wright, J. (1983). Children's processing of television: The informative functions of formal features. In J. Bryant and D. Anderson (Eds.), *Children's understanding of television* (pp. 35–65). New York: Academic Press

Jones T. (2011, April, 7). Designing for second screens: The Autumnwatch Companion. BBC. http://www.bbc.co.uk/blogs/researchanddevelopment/2011/04/the-autumnwatch-companion---de.shtml Accessed June 17, 2015.

Klein, J., Freeman, J., Harding, D., and Teffahi A. (2014). Assessing the impact of second screen. *Ofcom*.

Lochrie, M., & Coulton, P. (2012). Sharing the Viewing Experience Through Second Screens. *Proceedings of EuroiTV '12 ACM*.

Moulding, J. (2012) BBC Reveals Companion Screen App Strategy. http://www.v-net.tv/bbc-reveals-companion-screen-app-strategy. Accessed June 17, 2015.

Murray, J., Goldenberg, S., Agarwal, K., Chakravorty, T., Cutrell, J., Doris-Down, A., and Kothandaraman, H. (2012). Story-map: iPad companion for long form TV narratives. *Proceedings of EuroiTV '12 ACM*.

Nandakumar, A., & Murray, J. (2014). Companion apps for long arc TV series: supporting new viewers in complex storyworlds with tightly synchronized context-sensitive annotations. *Proceedings of TVX'14 ACM*.

Neate, T., Jones, M., & Evans, M. (2015, April). Mediating Attention for Second Screen Companion Content. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 3103-3106). ACM.

Nesbit, J., and Adesope, O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of educational research*, *76*(3), 413-448.

Nesbit, J., and Adesope, O. (2011). Learning from animated concept maps with concurrent audio narration. *The journal of experimental education*, 79, 209–230

Novak, J. (1990). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching*, 27(10), 937-949

Park, J., Blythe, M., Monk, A., & Grayson, D. (2006). Sharable Digital TV: Relating Ethnography to Design Through Un-useless Product Suggestions. *Proceedings of CHI '06 Extended Abstracts ACM*.

Robertson, S., Wharton, C., Ashworth, C., & Franzke, M. (1996). Dual Device User Interface Design: PDAs and Interactive Television. *Proceedings of the CHI'96 ACM*.

Van Dijk, J., and Vos, L. (2001). Searching for the holy grail. *New media and society*, 4(3), 443-465

Vatavu, R.-D., & Mancas, M. (2014). Visual Attention Measures for Multi-screen TV. *Proceedings of TVX'14 ACM*.

Appendix 1. Participants' gaze shifts cued by **auditory content changes**

Non-verbal to Narration

- 1%
- 14.5%
- 84.5%

Non-verbal to Interviewee

- 1%
- 5.5%
- 93.5%

Narration to Interviewee

- 0%
- 11%
- 89%

Narration to Non-verbal

- 1%
- 36.5%
- 62.5%

Interviewee to Narration

- 0%
- 0%
- 100%

Interviewee to Non-verbal

- 0%
- 52%
- 48%

*Narration to Interviewee: 9 participants   *Interviewee to Narration: 4 participants

- CompanionMap -> TV
- No shift from CompanionMap
- CompanionMap -> Off-screen

Appendix 2. Participants' gaze shift cued by **visual content change**

Filler to Portrait

2%
54.5%
44.5%

Filler to Illustration

2%
25%
73%

Illustration to Portrait

2%
43%
55%

Illustration to Filler

3%
23%
74%

Portrait to Illustration

2%
49%
49%

Portrait to Filler

0%
42%
58%

- TV-> TV
- TV -> CompanionMap
- TV-> Off-screen