INTERACTION WITH TELEVISION COMPANION APPS: FOUR FINDINGS AND A MODEL

John Dowell and Hana Kim
University College London

Companion apps for television programmes provide additional, synchronized and interactive content on mobile devices such as tablets or smartphones. With the television screen they create dual screen interfaces with multiple modalities and require viewers to actively manage their visual attention. We outline a model of interactions with companion apps used with information-rich television programmes where a primary purpose of the app is to support understanding and learning. Our model summarises perceptual and cognitive processes involved by drawing on theories and findings from Human Factors and the learning sciences. We use the model to assess CompanionMap, an app for accompanying science documentaries with synchronized, animated concept maps. We show how the model provides explanations for four findings obtained from an experiment with CompanionMap, including how users managed their visual attention and how they learnt about astrophysics when using the app to watch an astronomy programme.

INTRODUCTION/THESIS

Science programmes on television are the primary means of increasing public understanding of science (Wellcome Trust, 2013); they also make some of the greatest intellectual demands of viewers. Much of the art of the producer lies in creating TV science programmes that a diverse audience is able to engage with and learn from.

Science programmes are also a rich potential application for companion apps. Companion apps on web-enabled tablets and smartphones give their users access to content synchronized with the programme and can make TV programmes more compelling. Companion apps have been demonstrated with several genres of programme including fact-based programmes such as natural history programmes, but not with science programmes and not for augmenting the viewer's understanding, learning and engagement. The prospect of using companion apps for science programmes raises the question of in what form and with what kinds of interaction could a second screen augment the viewer's understanding and learning?

CompanionMap is a prototype companion app that uses animated concept maps to augment viewing of information-rich programmes such as science documentaries. In synchrony with a programme, new nodes appear on the concept map with their links joining to other nodes. The design of this app is encouraged by previous research into concept maps that finds a consistent benefit for learning outcomes and that animated versions are additionally effective (Nesbit &Adesope, 2006). It has also been shown that concept maps enhance learning when used in combination with a spoken recording (Adesope & Nesbit, 2013), although the same benefit could not be assumed when a concept map was used with a television programme competing for visual attention.

Dowell et al (2015) report an experiment to investigate whether CompanionMap improves understanding of a programme and learning, whether these improve with more interaction with the app, and whether users are able to manage their visual attention effectively over both screens. Using CompanionMap, participants watched two sections of a recorded astronomy documentary. Understanding and recall of the documentary was assessed with probe questions, eye gaze direction was assessed in relation to a content analysis of the programme, and subjective measures were obtained by questionnaires and interviews. Hence the experiment examined whether learning gains could be found when using CompanionMap and whether users would be able to successfully manage their visual attention over both screens. We summarise our findings for both of these questions in this paper.

If they are to be useful and usable, designs of companion apps for information-rich programmes need to be informed by research. CompanionMap was essentially the product of pure invention, even acknowledging the prior work on learning from animated concept maps. We have subsequently developed a theoretically-informed model of interactions with companion apps which we present in the following section. We show how the model provides explanations for key findings from the experiment with CompanionMap.

REVIEW: MODEL OF INTERACTIONS WITH COMPANION APPS

The model is for analysing interactions with companion apps specifically designed to support information-rich television programmes. The model characterizes those interactions in terms of three levels, each resting on a distinct cognitive theory: 'Attention' is based on 'multiple resource theory' (Wickens, 1992); 'resource and workload' is based on 'dual coding theory' (Clark & Paivio, 1991) and 'chunking' (Gobet et al, 2001), and; 'schema interaction' (Bransford et al, 1999; Rumelhart & Norman, 1976), and a cognitive theory of multimedia learning (Mayer, 2005). Each level affects the other and each implies design requirements for companion apps.

Attention that is interleaved and selective

Only the perception level of multiple resource theory (MRT) is recruited to this model. MRT distinguishes the modalities and codes of perception, each having dichotomous dimensions: visual and auditory (modalities), and; spatial and verbal (codes). Perceptual processing involves one or more of the possible four 'modality+code' channels: Visual-Spatial VS (e.g., analogue picture); Visual-Verbal VV (e.g., visual text); Auditory-Spatial AS (e.g., sound localisation and pitch), and; Auditory-Verbal AV (e.g., speech). MRT further assumes that perceptual processing allocates cognitive resources to a particular modality/code channel and that some modalities and codes have separate allocations of resources. Perception that involves channel combinations with the same modality or code will draw on shared resources and will compete for those resources. Watching TV uses VS and AV perceptual channels which draw on separate resource pools and so don't interfere with each other, at least until visual text appears on the screen, for example, to reinforce spoken information on news programmes. By adding one or more perceptual channels to watching TV, companion apps increase the demand for perceptual resources and the potential for interference with the perceptual channels used for watching the television programme.

Because companion apps typically run on a portable device, viewers must manage their visual attention across both screens. A strategy of 'Interleaved Selective Attention' (ISA) is necessary to look at the companion app while continuing to watch the TV programme. Significantly, attention is being managed over separate information sources but for the same task of watching and understanding an information-rich programme. The strategy is equivalent to, but significantly different from, interleaved multi-tasking strategies (de Fockert et al, 2001) when two or more tasks are being monitored and executed (Payne et al, 2007).

Companion apps for information-rich television programmes are primarily to support the assimilation of the programme content and the visual focus will necessarily be on the television screen. Attention to the companion app will necessarily be subsidiary to the television. Viewers will consciously shift their visual attention between the two screens and this distribution of visual attention will be significantly driven top-down by the schema the viewer constructs, moderated by the workload the viewer experiences which they will seek to maintain within comfortable limits. The direction of gaze is the overt expression of attention which is controlled by the central executive of working memory (Baddeley & Hitch, 1974).

Modalities relate to sense receptors (Moreno & Mayer, 2007) and ISA operates over hearing and vision. The auditory content of television programmes contains both verbally and spatially coded information; viewers of educational programmes such as science documentaries are believed to focus on auditory information more than visual information (Basil, 1994). Companion apps compete with the television screen for visual attention, but not for auditory attention since they will usually be silent. Companion app and television screen will not be within the same field of view and so visual attention of each must be interleaved. During visually un-informative parts of the TV programme, viewers will be most likely to move their focus of visual attention to the companion app under the direction of the schema they construct for the programme.

The visual content of TV programmes will usually be image based only whereas the visual content of companion apps can make use of both image and text. So TV visual content is a VS visual-spatial channel whilst the companion app visual content can be both VS and VV if visual text is used. Figure 1 graphically depicts the possible combination of verbal and spatial codings for the companion app. The VV channel of the companion app will compete for resources with spoken content of the television programme, an AV channel. For successful ISA over both verbal channels, it will be necessary for the visual text content of the companion app to be well coordinated with the spoken content of the television programme.

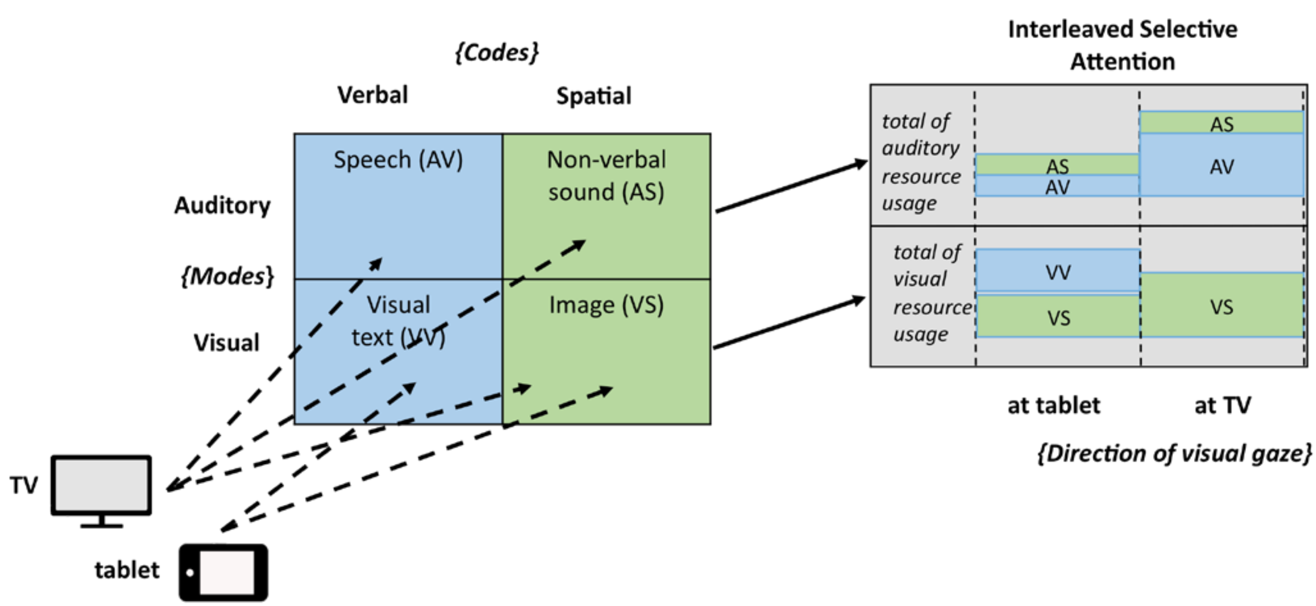

Figure 1: Schematic of attention management in interaction with a companion app.

To minimise the competition for verbal resources, visual text on the companion app should be simple, short and well synchronized with the spoken content of the television programme. Spatially coded information on the companion app will not interfere with spoken content of the television programme but will compete for resources with the visual imagery in the programme.

Working memory modes, codes and chunks

In principle, a companion app need not interfere with the viewer's perceptual processing of the television programme; in practice, companion apps will interfere unless they are compatible with the characteristics of working memory. The model identifies those characteristics; in essence, it is through their promotion of 'dual-codes' and 'chunking' that companion apps benefit working memory.

Working memory has both visual and spatial processing capabilities supervised by a central executive system (Baddeley & Hitch, 1974). Both verbal and spatial systems have non-persistent, temporary stores with limited capacity. Storage demands on working memory will deplete the resources available for processing (Just & Carpenter, 1992). The benefits of companion apps for information-rich programme viewing can be understood in relation to these working memory systems and their capacities.

Dual codes. First, the tendency for viewers to focus on the auditory information more when watching television programmes makes disproportionate demands on their verbal processing. By augmenting the verbal information in the programme with additional spatial information on the companion device, demand on verbal processing can be relieved.

Second, using two codes has been shown to be more effective than just one. Dual coding theory emphasizes "the importance of verbal associative structures and their spatial representation" (Clark & Paivio, 1991). Because of 'redintegration' (i.e., particular cues easily re-activate an entire representation), using dual codes facilitate remembering and retrieving information. Further, the central executive will be able to more effectively coordinate the two systems, leading not only to stabilising the information but also to decreasing the workload of working memory.

Summary chunking. Chunking is treating strongly associated elements as a group, while elements with low association become other chunks (Gobet et al, 2001). The demonstrable benefits of chunking for perception and for memory have encouraged the application of chunking to educational design.

Information-rich television programmes have the potential to overwhelm the viewers' processing resources which chunking can ameliorate. In contrast with its conventional usage which simply concatenates information, we propose 'summary chunking' in which information is both combined and abstracted. Summary chunks will relieve the demands on perception, freeing more resources in working memory and allowing more information to be stored.

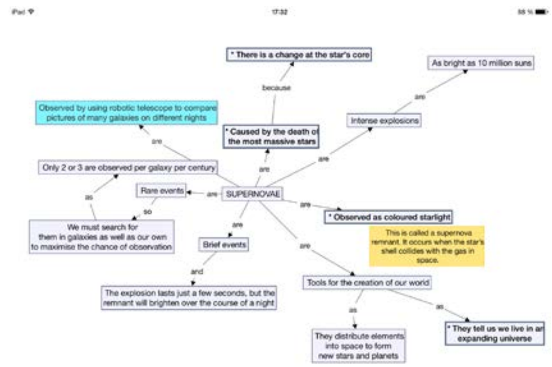

Figure 2. CompanionMap displaying an animated concept map about supernovae, synchronized with the astronomy programme 'The Seven Ages of Starlight' (BBC).

Guidelines for the design of companion apps to take advantage of the dual-code effect and the chunking effect are implied by this model:

i. The content of the companion app should be clear, brief and simple.
ii. Text and visual images should be used in combination.
iii. Text content should be combined where possible into abstracted chunks.

Schema interaction

Television programmes interact with viewers' knowledge represented in the form of schemata. A schema is the "primary meaning and processing unit of the human information processing system" (Rumelhart & Norman, 1976). Related schemata are joined in a network, an interrelated knowledge structure. Watching an information-rich programme is a form of learning in which the viewer's schemata change. They are extended by accretion, revised by tuning and re-structured when confronted by new information that does not fit existing schemata.

In watching a science documentary or other information-rich programme, viewers retrieve an existing relevant schema. Information presented in the programme that doesn't already exist in the schema but fits the structure will be added, a process of accretion. Some new information presented in the programme will cause the viewer's schema to be revised, for example, by dividing an existing concept into two new distinct concepts, a process of tuning. Some programmes will present information which is so incompatible with an existing schema that a new schema must created to accommodate it, replacing any previous related schema. Metacognitive experience will judge which information the viewer knows (Flavell, 1979) and will control schema learning to resolve the experience of not understanding the programme.

Mayer describes five cognitive theories of multimedia learning (Mayer, 2005) which all apply to using companion apps to watch information-rich programmes. The theories are: 1) selecting relevant words for processing in verbal working memory; 2) selecting relevant images for processing in visual working memory; 3) organising selected words into a verbal model; 4) organising selected images into a pictorial model; 5) integrating the verbal and pictorial representation with each other and with prior knowledge. The last of these theories applies to the level of schema learning.

NEW CONTRIBUTION: INTERPRETING INTERACTIONS WITH COMPANIONMAP

The effect of its concept map representation on participants' recall and understanding of a science documentary was investigated in an experiment with CompanionMap. The study also investigated participants' management of their visual attention over both screens, and their subjective experience of using the app. The study was conducted as a comparison of watching passages of the same astronomy programme with and without CompanionMap. The study of CompanionMap is reported comprehensively in Dowell et al (2015); here we select four of its findings and use the model to generate an explanation for each.

As a preliminary to that discussion, the model allows us to critically assess design features of CompanionMap. First, the short, clear, and simple labels used in the concept map conform with the model's prescription that visual verbal information on companion apps minimizes interference with the programme. Second, the verbal and spatial content of concept maps conform with the model's prescription that companion apps should promote dual-codes. Finally, the nodes of the concept map abstract the main concepts of the programme, conforming with the model's prescription that apps should promote summarized chunks. The four findings from the study are now discussed.

Eye gaze durations - participants' attention patterns

Participants' visual attention was investigated by recording durations of eye gaze on each screen and the frequency of eye gaze transitions between screens. The mean duration of 'looks' at the TV was 7.04 seconds (SD=3.91) and 3.15 seconds (SD=0.85) at CompanionMap. The mean frequency of gaze shifts was 75 (n.b., the programme section being viewed was some 6 minutes long). So participants' visual attention was marked by a high number of short looks with frequent shifts from one screen to the other, the television screen receiving twice as much visual attention as the companion app.

Interpretation. Participants' visual attention was governed by the Interleaved Selective Attention (ISA) strategy, both screens serving a common task with the TV screen as primary and the companion as auxiliary. Participants were integrating the abstracted concept labels from the concept map with the programme content, necessitating frequent short looks at the app. They were able to read the concept map while continuously listening to the programme soundtrack and integrate the labels on the map with their comprehension of the programme.

Gaze transitions with auditory cues

The experiment examined how visual attention switched from CompanionMap back to the TV screen in response to changes in the auditory content of the programme. The auditory content was categorised as either: non-verbal, narrator's voice, or interviewee's voice. Eye gaze transitions from CompanionMap to TV were analysed in relation to these cues. Participants were found to shift their gaze from the tablet to the TV 85.6% of times when the documentary changed from non-verbal to narration, 91.2% of times when non-verbal changed to interviewee, and 83.3% of times when narration changed to interviewee, and 100% of times when interviewee changed to narration. The striking finding is that people will look back at the TV screen when a voice starts speaking, be it a speaker speaking after a period of non-speech, or a change of speaker.

Interpretation. The model characterizes viewers as attending to both visual and auditory information when watching information-rich programmes and the auditory information dominating their attention. Hence our participants would usually switch their gaze back to the TV in response to a change in speaker regardless of what they were looking at on the tablet at the time. In watching information-rich programmes, verbal coding of information is primary. Reading the concept map means dividing their resources for verbally coded information, but the TV audio will be given priority. So when participants detect 'important new information going on', they will rapidly revert to the TV not in fact to see the narrator's or interviewee's face, but to no longer be seeing the map and to concentrate their verbal resources on the auditory speech.

Gaze transitions with visual cues

Transitions of visual attention from the TV screen to CompanionMap were also investigated in relation to changes in the visual content of the science documentary. The visual content was categorized as: filler (i.e. un-informative visuals); illustration, or; portrait (the interviewee's or narrator's face). Significant correlations with participants' eye gaze transitions from TV to CompanionMap were found: when the documentary content changed from filler to illustration, eye gaze shifted from TV to tablet on 78.7% of instances, and from illustration to filler on 77.6% of instances. Transitions in response to other cues were at chance level only.

Interpretation. Unlike the auditorally-cued transitions of eye gaze, the visually-cued transitions showed fewer clear patterns. This would be consistent with the model's characterization of attention favouring the auditory information and verbal information. Unlike auditory attention, visual attention could be selected rather than divided. The finding that participants would look away from the TV when illustrations changed to filler is consistent with the model's characterization of an effective management of visual attention, participants choosing to make best use of their visual attention resources and reduce the interference created by filler images. The significant tendency to look away from the TV screen just at the point that meaningless filler turned to meaningful illustration is particularly intriguing. The model characterizes this behaviour as viewers seeking a verbal label for the spatially-coded information, to help them interpret it and fit it into their developing schema.

Participants' learning with CompanionMap

The effect of CompanionMap on participants' understanding and recall of the programme was assessed using a set of probe questions. Participants watched one part of the programme using the CompanionMap and the other part without it. Questions probed knowledge of the programme that had been represented on the concept map, and knowledge of the programme that had not been on the concept map. A paired t-test indicated that participants' learning was higher when they used CompanionMap $T(15)=2.2$, $p=0.044$.

Interpretation. The benefit of CompanionMap for understanding and recalling the science documentary is due to its promoting dual-coding and chunking, and its direct cueing of schema.

Concept-maps are good for integrating related concepts (Nesbit &Adesope, 2006). When people look intermittently at the concept map on CompanionMap, it helps integrate all information from different codes. So this effect on dual-codes makes better use of the resources of working memory. The summarization chunks presented by the concept map reduce the demands on working memory. By labeling summarization concepts, the concept map provides a concrete text that enhances recall over abstract texts (Corkill et al, 1988). This chunking not only lessens the demands on working memory but also the demands on verbal capacities (Just & Carpenter, 1992). Workload is decreased because hierarchical relationships between concepts represented explicitly in visuospatial form augment semantic processing (Adesope & Nesbit, 2013).

CompanionMap's concept maps cue retrieval of relevant schema, its animated nodes cue the process of accretion of information from the TV programme into the schema, and the relationships between nodes cue the processing of relationships within the schema. The concept map supports tuning of schema by explicitly representing concepts and relationships from the programme that necessitate revisions to existing contents of the schema and extensions to the structure of the schema.

DISCUSSION

A model of interaction with companion apps supporting information-rich television programmes is proposed in this paper. We have shown that the model is able to explain key findings from the evaluation of CompanionMap, a prototype companion app for science documentaries. The findings related to the pattern of gaze shifts between companion app and television screen, the cues that affect when gaze shifts occur from one screen to another, and the learning benefits of using the app. Each finding was explained using the model. ISA was applied to explain eye gaze duration, and eye gaze transition, and dual-coding, summary chunking, and schema learning were used to explain participants' learning. The model characterises core cognitive processes occurring when a companion app is used to enhance watching TV programmes and our study indicates that it is able to inform the design of more effective companion apps.

REFERENCES

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

Baddeley, A. & Hitch, G. (1974). Working Memory. *Psychology of learning and motivation*, 8, 47–89.

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

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.

Bransford, J. D., Brown, A. L. & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school*. National Academy Press.

Clark, J.M. & Paivio, A. (1991). Dual coding theory and education. *Educational psychology review*, 3(3), 149-210.

Corkill, A. J., Glover, J. A. & Bruning, R. H. (1988). Advance organizers: Concrete versus abstract. *Journal of Educational Research*, 76-81.

de Fockert, J. W., Rees, G., Frith, C. D., & Lavie, N. (2001). The role of working memory in visual selective attention, *Science*, 1803-1806.

Dowell, J., Malacria, S., Kim, H. & Anstead, A. (2015) Companion apps for television programmes: representation and interaction. *Personal and Ubiquitous Computing* (in press).

Flavell, J. H., (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist*, 34(10), 906-911.

Gobet, F., Lane, P., Croker, S., Cheng, P., Jones, G., Oliver, I, & Pine, J. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Science*, 5(6), 236-243.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological review*, 99, 1, 122-149.

Logie, R. H. (2011). The functional organization and capacity limits of working memory. *Current Directions in Psychological Science*, *20*(4), 240-245.

Mayer, R. (2005). Cognitive theory of multimedia learning. In Mayer, R. (ed.) *The Cambridge handbook of multimedia learning*, CUP.

Moreno, R. & Mayer, R. (2007). Interactive multimodal learning environments. *Educational psychology review*, Springer (2007), 309–326.

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

Payne, S. J., Duggan, G. B. & Neth, H. (2007). Discretionary task interleaving: heuristics for time allocation in cognitive foraging. *Journal of Experimental Psychology*, 136(3), 370.

Rumelhart, D. E. and Norman, D. A. (1976). *Accretion, tuning and restructuring: Three modes of learning*, La Jolla, CA: Center for Human Information Processing, University of California, San Diego.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design, *Learning and instruction*, 4(4), 295-312.

Wickens, C. D. (1992). *Engineering psychology and human performance*. HarperCollins Publishers, New York, NY.

Wellcome Trust (2013). *Engaging with Science*, Wellcome Trust Monitor Wave 2 (May, 2013), www.wellcome.ac.uk/monitor.