Title
Does Video Content Facilitate or Impair Comprehension of Documentaries?The Effect of Cognitive Abilities and Eye Movement Strategy

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 41(0)

Authors
Zheng, Yueyuan
Ye, Xinchen
Hsiao, Janet H.

Publication Date
2019

Peer reviewed

Does Video Content Facilitate or Impair Comprehension of Documentaries? The Effect of Cognitive Abilities and Eye Movement Strategy

Yueyuan Zheng (u3514160@connect.hku.hk)
Department of Psychology, University of Hong Kong,
Pokfulam Road, Hong Kong

Xinchen Ye (yexc@connect.hku.hk)
Department of Psychology, University of Hong Kong,
Pokfulam Road, Hong Kong

Janet H. Hsiao (jhsiao@hku.hk)
Department of Psychology, University of Hong Kong,
Pokfulam Road, Hong Kong

Abstract

It remains unclear whether multimedia facilitates or impairs knowledge acquisition. Here we examined whether subtitles and video content facilitate comprehension of documentaries consisting of statements of facts and whether the comprehension depends on participants' cognitive abilities and eye movement strategies during video watching. We found that subtitles facilitated comprehension regardless of participants' cognitive abilities or eye movement strategies for video watching. In contrast, with video content but not subtitles, comprehension depended on participants' auditory working memory, task switching ability, and eye movement strategy. Through the Eye Movement analysis with Hidden Markov Models (EMHMM) method, we found that a centralized (looking mainly at the screen center) eye movement strategy predicted better comprehension as opposed to a distributed strategy (with distributed regions of interest) after contributions from cognitive abilities were controlled. Thus, whether video content facilitates comprehension of documentaries depends on the viewers' eye movement strategy in addition to cognitive abilities.

Keywords: multimedia; eye-movement; hidden Markov model

Introduction

Multimedia learning refers to knowledge construction from both verbal and pictorial information, with the verbal form including spoken words or printed texts, and the pictorial form including illustrations, graphs, pictures, photos, animations, and videos (Mayer, 2014a). The modality principle suggests that it is generally beneficial to receive both visual and audio information in the learning process (Low & Sweller, 2014). Similarly, subtitles have been shown to enhance learning: same-language subtitles in video advertisements were shown to enhance the viewers' memory of the brand and slogan (Brasel & Gips, 2014), and watching recorded lectures with subtitles was associated with better comprehension performance (Kruger & Steyn, 2014).

The cognitive theory of multimedia learning further posits that learning would be undermined if multiple sources of information are received from the same perceptual channel, and would be facilitated if different sources of information are received from independent channels (Mayer, 2014b). For example, people displayed worse learning outcome when watching animations with on-screen texts than without because both text and animation information came from the visual modality (Mayer, Heiser & Lonn, 2001). However, some more recent research reported no trade-off between image and text processing (Perego, Del Missier, Porta & Mosconi, 2010; Kruger, Soto-Sanfiel & Doherty, 2017), and that participants learning through multimedia displayed better knowledge acquisition and improved content comprehension as compared with those who learned through text reading or traditional lectures in academic learning (Starbek, Erjavec, Starcic & Peklaj, 2010).

The inconsistent findings in the literature may be due to differences in the type of learning materials used and learners' language proficiency across studies. The effect of multimedia learning may depend on how the most important information for comprehension was delivered. For example, for documentaries containing mainly statements of facts, the auditory narratives may contain most of the information, and thus video content may be distracting and impair comprehension whereas same-language subtitles may be helpful especially for second-language learners. Indeed, unsubtitled videos were reported to create a higher cognitive load (as indicated by pupil diameter change) and frustration levels (as measured by EEG) than subtitled versions for students learning English as a second language (Kruger, Hefer & Matthew, 2013). In contrast, for learning involving graphic demonstrations, video content may contain additional information, which may compete with on-screen texts for cognitive resources.

In addition, individual differences in cognitive abilities and strategies may also influence whether multimedia helps or impairs learning. For example, the ability to flexibly switch attention among various sources of information and to focus on the relevant information while inhibiting irrelevant information may be important for successful multimedia learning (Miyake et al., 2000). Indeed, research has reported that readers with high working memory capacity were more effective in selecting and integrating information and achieved better comprehension in multimedia learning (Schnotz, 2005; Fenesi, Kramer & Kim, 2016). Similar results were found for those with better task switching ability (Baadte, Rasch & Honstein, 2015).

Another possible factor is individual differences in cognitive strategy or perception style, which may be better revealed through eye tracking (Hyona, 2010; van Gog & Scheiter, 2010). Previous research using eye tracking to understand multimedia processing typically only focused on group level comparisons, such as comparing adults' and children's learning (D'Ydewalle & De Bruycker, 2007). However, recent studies have reported significant individual differences in eye movement patterns that can reflect differences in cognitive strategy and perception style (e.g., Chan, Chan, Lee & Hsiao, 2018). It is possible that participants adopting different eye movement strategies during multimedia learning differed in whether multimedia facilitates or impairs learning.

Here we aimed to examine how individual differences in cognitive abilities and eye movement strategies modulate multimedia learning of documentaries consisting of mainly statements of facts, and thus the important information would be mainly in the auditory narratives. Specifically, we examined how adding subtitles and video content would influence participants' comprehension of auditory science documentaries, and whether participants' working memory capacity, switching ability, and eye movement strategy for video watching could predict the comprehension of the documentaries. We recruited native speakers of the language used in the documentaries to control for language experience. To discover common eye movement strategies for video watching from the participants and quantitively measure individual differences in eye movement pattern, we used the Eye Movement analysis with Hidden Markov Models (EMHMM, Chuk, Chan, & Hsiao, 2014) method to analyze eye movement data. In this method, each participant's eye movement pattern during a visual task is summarized using an HMM, including personalized regions of interest (ROIs) and transition probabilities among the ROIs. Individual HMMs then can be clustered according to similarities to discover common strategies in the participants (Coviello, Chan & Lanckriet, 2014). The similarity between a participant's eye movement pattern to a common strategy can be assessed as the log-likelihood of the participant's eye movement data being generated by the HMM of the common strategy. This quantitative measure of eye movement pattern similarity then can be used to examine the associations between eye movement patterns and other cognitive measures. We hypothesized that (1) subtitles are helpful in the comprehension of the documentaries because of the exact match to the content of the documentaries, which may facilitate retrieving the meanings of the auditory narratives, (2) people with higher working memory capacity or task switching ability may show better comprehension when learning from documentaries with video content, and (3) people with more explorative eye movement patterns during video watching may be more distracted by video content, leading to worse comprehension performance.

Methods

Participants

Sixty native Mandarin speakers (40 females, 18 to 30 years old, M = 21.07, SD = 3.32) were recruited from the University of Hong Kong. Participants were from different majors except for ecology, astronomy, geography and chemistry, which were the topics of the documentary clips used here. All participants reported normal or corrected-to-normal vision.

Materials

The materials consisted of 16 documentary video clips in ecology, astronomy, geography, and chemistry, with 4 clips in each topic. The length of each clip was 75 s, and the resolution was 1920 x 1280 pixels. All clips used were produced by China Central Television (CCTV) and Shanghai Education Television (SETV) and were accessible to the public. The clips were selected based on the following criteria: a) Mandarin narratives; b) with simplified Chinese subtitles; c) not translated from foreign languages; d) produced as statements of facts, as to ensure the understandability of the clips to native speakers and to avoid possible linguistic biases.

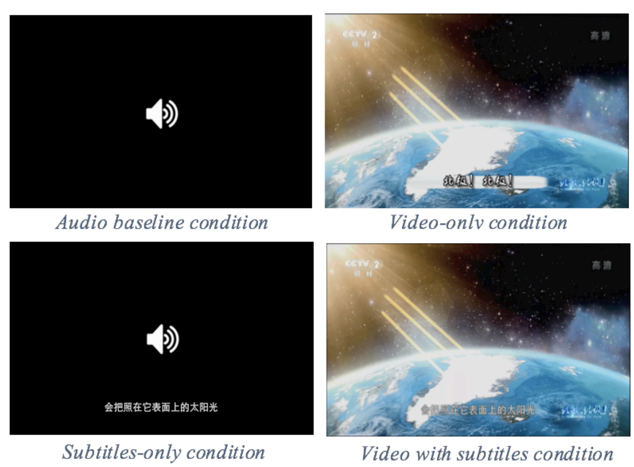


Figure 1: Four conditions of documentary presentation.

Each documentary could be presented in 4 different conditions: a) *Audio baseline* condition (i.e., narrated sound only without video content or subtitle, with a static icon presented at the center of a black screen); b) *Subtitles-only* condition (i.e., Audio baseline condition with original subtitles); c) *Video-only* condition (i.e., narrated sound with full-screen video content, with a fixed title masking the original subtitles); and d) *Video with subtitles* condition (i.e., narrated sound with both video content and original subtitles; Figure 1). Among the 16 original video clips, half of them had subtitles located on the bottom left of the screen, while the other half set the subtitles on the bottom center.

Design

Here we examined how video content and subtitle affected comprehension of documentaries. The independent variables were video content (with vs. without) and subtitle (with or without), resulting in four experimental conditions: audio baseline, subtitles only, video only, and video with subtitles. Each participant viewed 16 clips in total with 4 clips in each condition (one from each topic); the clips used in the 4 conditions were counterbalanced across participants. The dependent variable was accuracy in answering comprehension questions related to the clips. Repeated measures ANOVA was used for the data analysis.

In a separate analysis, we examined whether eye movement strategies used in the video only and video with subtitles conditions could predict comprehension of documentaries. The EMHMM approach was used (Chuk et al., 2014). More specifically, for each of the two video conditions separately, we used one HMM to summarize a participant's eye movement pattern across all clips. We then clustered the individual HMMs into two groups according to similarities to discover two representative eye movement strategies among the participants. Participants' eye movement pattern similarity to the two strategies then could be quantitatively assessed by calculating the log-likelihood of their data being generated by the HMM of the representative strategies. We examined whether this eye movement pattern similarity measure could predict participants' comprehension.

We also examined whether participants' cognitive abilities could predict comprehension of documentaries. Participants completed a n-back task for testing working memory capacity (Owen, McMillan, Laird & Bullmore, 2005), a Tower of London task for assessing executive function and planning abilities (Shallice, 1982), and a multitasking task for testing task switching abilities (Pashler, 2000). We then performed a hierarchical analysis to examine whether eye movement pattern could still predict comprehension after variation due to cognitive abilities was controlled.

Procedure

Comprehension of Documentaries During the task, participants' eye movements were recorded by an eye tracker, Eyelink1000. The tracking mode was pupil and corneal reflection with a sampling rate of 1000 Hz. Stimuli were displayed on a 22" CRT monitor with a resolution of 1920 by 1440 pixels and 150 Hz frame rate. The viewing distance was 60 cm. The standard nine-point calibration procedure was carried out before the experiment and whenever the drift correction error was larger than 1° of visual angle. Each trial started with a white solid dot appearing at the center of the screen. Participants were asked to look at the dot whenever it appeared for drift correction. Afterwards, a documentary clip was played in full screen. After each clip, participants were asked to answer 6 aurally-presented multiple-choice questions (MCQs) according to the content. The MCQs were presented one at a time binaurally in Mandarin, and the voice was synthesized by the online Baidu voice producer. Participants could replay each question unlimited times before their response. Participants performed the task in the 4 different documentary presentation conditions in separate blocks, with the block order counterbalanced across participants. They proceeded to the cognitive tasks described below after the comprehension task.

Cognitive Tasks

1. N-back Test: Two-back tests with 3 types of stimuli, including visual English letters, spoken numbers, and irregular shapes (Figure 2A) were used to test visual and verbal working memory. For each type of stimuli, participants were presented with 30 items one at a time, each for 2.5 s with a 0.5 s interval (Lau et al., 2010), and asked to judge whether the item presented in a trial matched the one that appeared 2 trials back.

2. Tower of London Test: Participants were asked to move 3 color discs one at a time from an initial position to match a goal position with the minimum number of moves, and to plan the moves in mind before execution (Figure 2B; Phillips, Wynn, McPherson & Gilhooly, 2001). Participants completed 10 trials. The total number of moves, total execution time, and total preplanning time before executing the first move were measured. Five practice trials were provided.

3. Multitasking Test: Four types of figures with different combinations of shapes and fillings were used as the stimuli (Figure 2C). The stimuli were presented one at a time in either the top or the bottom half of a box at the center of the screen (Figure 2C, left). Participants were asked to perform a dual task where they judged the shape of the figure (the shape task) as fast and correctly as possible when the figure was presented in the top half of the box, and judged the number of dots in the filling of the figure (the filling task) when the figure was presented in the bottom half of the box (Stoet, O'Connor, Conner & Laws, 2013). The figure was presented for 2500ms, followed by a 500ms blank screen. Participants were asked to respond by the end of the 3-second trial. A shape-only and a filling-only task were tested sequentially before the dual-task to measure participants' baseline behavior where no task switching was involved. Their task switching ability was measured as the response time (RT) in the dual task minus the average RT during the two no-switching tasks.

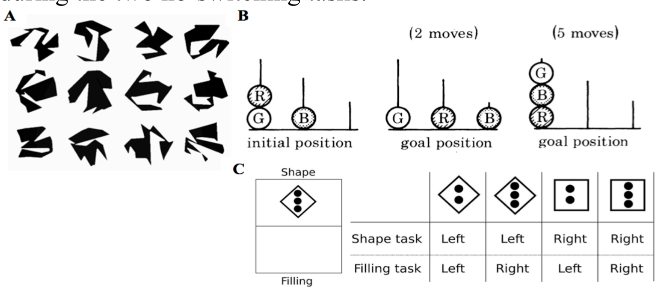


Figure 2: (A) 12 pictorial stimuli (Attneave and Arnoult structures) in the 2-back test, (B) Example of Tower of London test, (C) Stimuli used in the multitasking test.

Results

Effect of Video Content and Subtitle

The results showed a significant main effect of subtitle, $F(1, 59) = 13.359$, $p = .001$ (Figure 3). There was no main effect of video content or interaction between video content and subtitle. This result suggested that subtitles, but not video content, facilitated comprehension of documentaries.

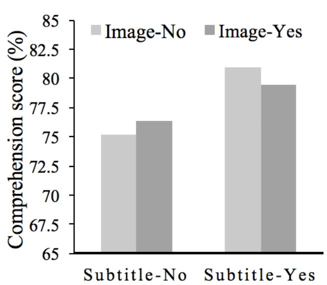
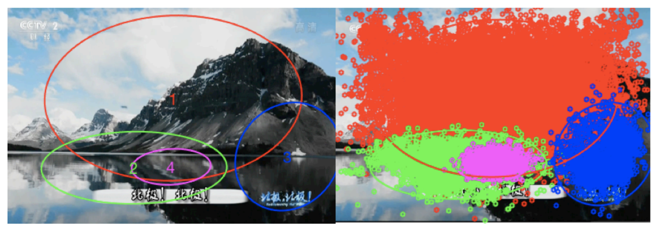


Figure 3. Effect of video content and subtitle on the comprehension of documentaries.

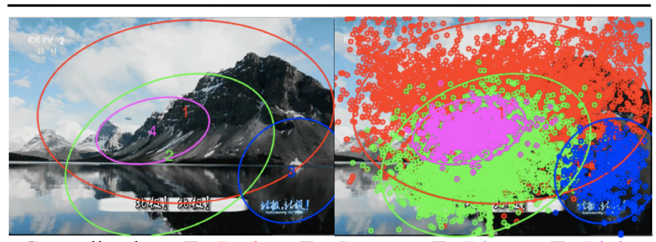
Eye Movement Strategies for Viewing Videos

Using the EMHMM method, for the video-only and video with subtitles conditions separately, we summarized each participant's eye movement pattern using an HMM. For each HMM, a variational Bayesian method was used to determine the optimal number of ROIs. More specifically, we ran each HMM with a different number of ROIs (ranging from 1 to 6) 300 times with a random initialization each time and selected the model with the largest log-likelihood given the data. We then clustered all individual HMMs into two groups and generated a representative HMM for each group with the number of ROIs set to 4.

Figure 4 shows the results of the *Video-only* condition. In the strategy on the top, after an initial fixation at the center of the video, participants had 8% of probability to look at either the blue ROI containing a logo on the bottom right or the green and the pink ROIs containing the fixed title on the bottom center of the screen. Afterwards, they tended to stay in the same ROI or switch back to the red ROI, or occasionally switched among the green, blue, and pink ROIs. We referred to this strategy as the distributed strategy. 46 participants were classified in this group. In contrast, in the strategy on the bottom, there were overlapping ROIs around the screen center, and participants mainly focused at the center of the screen. We referred to this strategy as the centralized strategy. There were 14 participants in this group.



Distributed	To Red	To Green	To Blue	To Pink
Priors	1.0	.00	.00	.00
From Red	.92	.05	.02	.01
From Green	.37	.56	.07	.00
From Blue	.38	.12	.47	.04
From Pink	.27	.00	.08	.64



Centralized	To Red	To Green	To Blue	To Pink
Priors	.46	.44	.01	.00
From Red	.91	.08	.01	.00
From Green	.18	.78	.04	.01
From Blue	.17	.39	.44	.00
From Pink	.03	.04	.00	.93

Figure 4. Distributed (top) and centralized strategies in the video only condition. Ellipses show ROIs as 2-D Gaussian emissions. The table shows transition probabilities among the ROIs. Priors show the probabilities that a fixation sequence starts from the ellipse. The image on the right shows raw eye fixation data and their ROI assignments.

To better understand the relationship between eye movement pattern and comprehension performance, following previous studies (e.g., Chan et al., 2018), we defined a Distributed-Centralized scale (D-C scale) for each participant as

$$Distributed\text{-}Centralized\ scale = \frac{D - C}{|D| + |C|}$$

Where D is the log-likelihood of the participant's eye movement data being generated by the representative HMM of the distributed strategy, and F is the log-likelihood of the participant's data being generated by the representative HMM of the centralized strategy. This log-likelihood measure reflects the similarity of the participant's eye movement pattern to the representative strategies. A more positive value in the D-C scale indicated higher similarity to the distributed strategy, whereas a more negative value indicated higher similarity to the centralized strategy. We found that participants' D-C scale was negatively correlated with comprehension performance in the video-only condition, r = -.291, p = .024: the more distributed the pattern, the lower the performance in comprehension (Figure 5A).

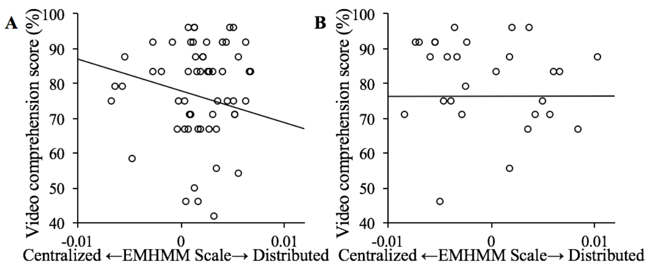
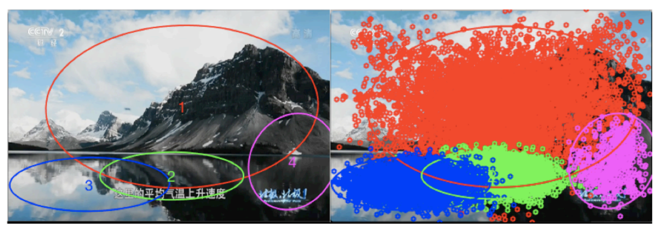


Figure 5. Correlation between eye-movement pattern and comprehension performance in the (A) video-only, and (B) video with subtitles condition.

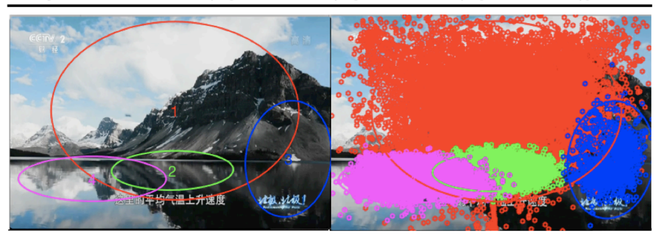
A similar analysis was conducted with eye movement data in the video with subtitles condition. Figure 6 shows the results of clustering participants' eye movement patterns into 2 groups. The 2 groups showed similar concentrations on the ROIs on the bottom left and bottom center of the screen, where the subtitles were located, in addition to the screen center. Group 1 strategy showed a higher probability to look at the subtitle regions after looking at the screen center. One-third of the participants (20 out of 60) adopted Group 1 strategy (one participant's eye movement data was invalid due to technical problems). We also measured participants' eye movement pattern similarity using the Group 1-2 scale in the same way as the D-C scale, and found that it did not correlate significantly with comprehension performance (Figure 5B).

Effect of Cognitive Abilities and Eye Movement Strategy on Comprehension

The above results suggested that in the video-only condition, participants' online eye movement pattern (D-C scale) predicted participants' comprehension. In addition to eye movement strategy, we found that comprehension performance was also significantly correlated with auditory working memory ability as measured in the n-back task, r = .337, p = .008, and task switching ability as measured in the multitasking test r = .262, p = .043. To examine whether eye movement pattern significantly contributed to comprehension after variation due to cognitive abilities was controlled, a three-stage hierarchical multiple regression was conducted to predict comprehension score. At stage one, auditory working memory capacity (N-back test) contributed significantly to the regression model, F(1,58) = 7.432, p = .008, and accounted for 11.4% of the variation. Adding task switching ability (Multitasking test) to the regression model explained an additional 5.5% of the variation in comprehension and the change in R^2 was significant, F(2,56) = 5.813, p = .005. Finally, introducing eye movement pattern (D-C scale) explained an additional 7.9% of the variation in comprehension score and this change in R^2 was significant, F (3,57) = 6.172, p = .001. Thus, when watching video documentaries without subtitles, participants' online eye movement behavior played an important role in comprehension in addition to cognitive abilities. A similar regression analysis was conducted for predicting comprehension performance in the video with subtitles condition, and no significant predicting variable was found.



Group1	To Red	To Green	To Blue	To Pink
Priors	0.98	.00	.02	.00
From Red	.74	.14	.11	.02
From Green	.22	.76	.01	.01
From Blue	.24	.01	.74	.01
From Pink	.23	.13	.13	.51



Group2	To Red	To Green	To Blue	To Pink
Priors	.99	.00	.00	.01
From Red	.79	.10	.01	.09
From Green	.22	.76	.01	.00
From Blue	.22	.13	.53	.11
From Pink	.24	.00	.01	.75

Figure 6. The two strategies observed in the video-with-subtitles condition.

Discussion

The present study aimed to investigate the role of video and subtitle in knowledge-based multimedia learning, with the effect of individual differences in multimedia processing including cognitive abilities and eye movement strategies considered. We hypothesized that for documentaries consisted of statements of facts, where auditory narratives provide most of the information, subtitles would facilitate comprehension due to the exact match to the content. In contrast, with video content, the comprehension may depend on participants' working memory, planning, and task switching abilities, as well as their eye movement strategy. Consistent with our hypothesis, we found that subtitles facilitated comprehension whereas video content did not. Through the EMHMM method, we discovered the distributed and centralized eye movement strategies in watching videos without subtitles. Interestingly, the more similar participants' eye movement pattern to the distributed strategy, the worse their comprehension in the video-only condition. Hierarchical regression analysis further showed that, while both auditory working memory and task switching abilities were significant predictors for comprehension, participants' eye movement pattern contributed significantly to comprehension after variation due to these cognitive abilities was controlled. This result showed that the facilitation of video content in the comprehension of documentaries depended on participants' online eye movement strategy in addition to working memory and task switching abilities. In contrast, participants' comprehension in the video with subtitles condition did not depend on either cognitive abilities or eye movement strategy.

Our results showed that adding subtitles is beneficial to knowledge acquisition of documentaries consisting of statements of facts. Previous studies on the effect of subtitles have reported inconsistent findings, with some showing that on-screen text is distracting (Mayer, Heiser & Lonn, 2001) whereas others suggesting facilitating effects (Starbek et al., 2010). This inconsistency may be due to differences in the amount of information carried in each medium during multimedia learning. In cases where pictorial stimuli contain important content for knowledge acquisition, simultaneous on-screen texts may be distracting (e.g., Mayer, Heiser & Lonn, 2001). In contrast, for materials where auditory narratives already provide most of the information for learning, such as the documentaries used in the current study, subtitles that match well with the auditory narratives may help maintaining participants' attention to the content of the knowledge and consequently facilitate comprehension (Kruger, Hefer & Matthew, 2013).

The finding that the distributed eye movement strategy, as opposed to the centralized strategy, was correlated with worse comprehension may be related to how attention was allocated in these two cases. According to the cognitive theory of multimedia learning (e.g., Mayer, 2014), engaging in active eye movement planning as demonstrated in the distributed strategy where there were specific ROIs located at different regions of the video may increase the cognitive load, resulting in decreased attentional resources for listening comprehension. The EMHMM method allows discovery of representative eye movement strategies from individual patterns in a data-driven fashion and provides a quantitative assessment of eye movement pattern similarities, leading this novel finding.

We also observed that in the video with subtitles condition, most participants focused on the subtitle locations in addition to the screen center, and participants' comprehension could not be predicted by either eye movement strategies discovered in this condition, working memory ability, or task switching ability, in contrast to the video-only condition. This result suggests that subtitles with video may help maintain participants' attention to a specific location of the video and reduce the possibility of active eye movement planning to other regions of the video, resulting in more attentional resources to listening comprehension and reduced task switching requirements. The redundancy effect from subtitles may also decrease the demands on working memory.

Previous research has suggested that cognitive abilities such as working memory capacity could modulate multimedia learning effects (Fenesi, Kramer & Kim, 2016). Here we further showed that in addition to working memory and task switching abilities, comprehension performance in the video-only condition could be predicted by online eye movement strategy: people who adopted a more centralized eye movement pattern had better comprehension. Future work will examine whether an explicit instruction of using a centralized eye movement strategy during video documentary viewing will facilitate comprehension.

To conclude, here we showed that for knowledge acquisition from auditory narratives, subtitles facilitated comprehension, whereas with video content, comprehension depended on participants' working memory and task switching abilities, as well as online eye movement strategy. When watching videos without subtitles, participants' comprehension could be facilitated by better auditory working memory and task switch abilities, and a more centralized eye movement pattern. In contrast, when watching videos with subtitles, subtitles seemed to have attracted and stabilized eye movements as well as reduced the demands on working memory and task switching, and thus neither cognitive abilities nor eye movement strategy could predict comprehension performance. These findings demonstrated the importance of taking individual differences into account in the research on instructional design and science of learning, and eye tracking with EMHMM provides a useful tool for revealing and quantitatively assessing these individual differences.

Acknowledgements

We are grateful to RGC of Hong Kong (Project # 17609117 to Hsiao).

References

Baadte, C., Rasch, T., & Honstein, H. (2015). Attention switching and multimedia learning: The impact of executive resources on the integrative comprehension of texts and pictures. *Scandinavian Journal of Educational Research*, *59*(4), 478-498.

Brasel, S. A., & Gips, J. (2014). Enhancing television advertising: Same-language subtitles can improve brand recall, verbal memory, and behavioral intent. *Journal of the Academy of Marketing Science*, *42*, 322-336.

Chan, C. Y. H., Chan, A. B., Lee, T. M. C., & Hsiao, J. H. (2018). Eye-movement patterns in face recognition are associated with cognitive decline in older adults. *Psychonomic Bulletin & Review*.

Chuk, T., Chan, A. B., & Hsiao, J. H. (2014). Understanding eye movements in face recognition using hidden Markov models. *Journal of Vision*, *14*(11), 8-8.

Coviello, E., Chan, A. B., & Lanckriet, G. R. (2014). Clustering hidden Markov models with variational HEM. *The Journal of Machine Learning Research*, *15*(1), 697-747.

D'Ydewalle, G., & De Bruycker, W. (2007). Eye movements of children and adults while reading television subtitles. *European Psychologist*, *12*(3), 196-205.

Fenesi, B., Kramer, E., & Kim, J. (2016). Split-attention and coherence principles in multimedia instruction can rescue performance for learners with lower working memory capacity. *Applied Cognitive Psychology, 30*(5), 691-699.

Hyona, J. (2010). The use of eye movements in the study of multimedia learning. *Learning and Instruction, 20*(2), 172-176.

Kruger, J. L., Doherty, S., & Soto-Sanfiel, M. T. (2017). Original language subtitles: their effects on the native and foreign viewer. *Comunicar: Media Education Research Journal*, *25*(50), 23-32.

Kruger, J. L., Hefer, E., & Matthew, G. (2013). Measuring the impact of subtitles on cognitive load: Eye tracking and dynamic audiovisual texts. In *Proceedings of the 2013 Conference on Eye Tracking South Africa* (pp. 62-66). ACM.

Kruger, J. L., & Steyn, F. (2014). Subtitles and eye tracking: reading and performance. *Reading Research Quarterly*, *49*, 105-120.

Lau, E. Y. Y., Eskes G. A., Morrison, D. L., Rajda, M., Spurr, K. F. (2010). Executive function in patients with obstructive sleep apnea treated with continuous positive airway pressure. *J. Int. Neuropsych. Soc., 16,* 1077- 1088.

Low, R., & Sweller, J. (2014). The modality principle in multimedia learning. In *The Cambridge Handbook of Multimedia Learning, Second Edition*, 227-246.

Mayer, R. E. (2014). Introduction to multimedia learning. In R. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning, Second Edition*, 1-24. Cambridge: Cambridge University Press.

Mayer, R. E. (2014). Cognitive theory of multimedia learning. In *The Cambridge Handbook of Multimedia Learning, Second Edition*, 43-71. Cambridge University Press.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, *93*(1), 187-198.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49-100.

Perego, E., Del Missier, F., Porta, M., & Mosconi, M. (2010). The cognitive effectiveness of subtitle processing. *Media Psychology*, *13*(3), 243-272.

Phillips, L. H., Wynn, V. E., McPherson, S., & Gilhooly, K. J. (2001). Mental planning and the Tower of London task. *Q. J. Exp. Psychol. - A*, 54, 579–597

Schnotz, W. (2005). Integrated model of text and picture comprehension. In R. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd ed.), 72–103. Cambridge: Cambridge University Press.

Starbek, P., Erjavec, M. Starcic, & Peklaj, C. (2010). Teaching genetics with multimedia results in better acquisition of knowledge and improvement in comprehension. *Journal of Computer Assisted Learning, 26*(3), 214-224.

Stoet, G., O'connor, D., Conner, M., & Laws, K. (2013). Are women better than men at multi-tasking? *BMC Psychology, 1*(1), BMC Psychology, 12/2013, Vol.1(1).

Van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction*, *20*(2), 95-99.