The effect of using a talking head in academic videos

An EEG study

Diana Díaz, Rafael Ramírez, Davinia Hernández-Leo

Department of Information Technologies

Universitat Pompeu Fabra

Barcelona, Spain

dianacarolinadiazcruz@yahoo.com, {rafael.ramirez, davinia.hernandez}@upf.edu

Abstract— This paper presents a study designed to understand the effect of using a talking head in academic videos frequently used in video-based learning approaches, such as Massive Open Online Courses. The experiment consisted of exposing participants to videos about different types of open software licenses. Each participant was exposed to 3 videos, each with a different condition: instructor always present (talking head condition), only instructor's voice present (only audio condition), and instructor presented only at the beginning of the video (mixed condition). Dependent variables included cognitive load, and emotional states (valence and arousal) obtained with electroencephalography, a personal assessment of the difficulty of the material, personal opinion regarding the social presence and performance in a memory test. The results indicate an increase in the cognitive load in the mixed condition, which may have implications regarding the use of the talking head in the design of academic videos.

Keywords—academic videos; design; talking head; MOOC; EEG

I. INTRODUCTION

Lately there has been a proliferative use of the talking head in the academic videos largely used in Massive Open Online Courses (MOOC). However, despite its use implies a high effort in production (in terms of time and economic resources) [1], there is still a lack of clarity about its effects for educational purposes.

Existing studies have explored the role of diverse design elements for academic videos in different contexts and subject matters. The study presented in this paper focuses on the appearance (or not-appearance) of a talking head as the design element under study. The context of the study is a higher engineering education setting, in which the content of the videos is around open software licenses. The experimental design includes a control of the content difficulty, the evaluation of the emotional levels, cognitive load, and performance. Instruments used include electroencephalography (EEG), questionnaires and a memory test.

The remainder of the paper is structured as follows: Section 2 describes related work. The materials and methods of the experiment are explained in section 3, while the results are presented in section 4. Section 5 discusses the main conclusions.

II. RELATED WORK

The use of the talking head in academic environments can be supported from several theories: a cognitive one implies considering the cognitive load theory, it suggests the existence of a mental burden that is imposed when a task is performed [2]. There are two types: “Intrinsic load” created by the inherent complexity of the content [2]; and “Germane load” or cognitive resources required to handle intrinsic cognitive load [3], that can be supported using the right educational tools. It is not clear yet how the talking head can contribute to the appearance of the second type of load. However, literature and research provide clues about the talking head use consequences. For example: working memory theories imply we have a limited capacity for visual and verbal information, therefore it has been suggested a benefit in the talking head use, due to the possibility it gives to keep information in a double channel [4]. Moreover, some theories show that placing words near corresponding graphics promotes learning [1], fact inherent in the talking head use. This can be associated to a decrease in load because attention does not have to split [4]. Nevertheless it is not clear if these benefits get lost considering that there is additional content in the presentation that can be attended.

Previous research using eye-tracker, showed that subjects spent a predominant time looking at the talking head over the rest of the content [1]. Notwithstanding the differences in the learning styles of the learners make difficult to find out if this is positive or negative (being a distracter) under the cognitive load scope. Consequently it has been found repeatedly clear benefits in the use of the lecture from an emotional approach. The main benefit is that it produces “social presence” that has been associated with student’s satisfaction [5].

Going deeper from the previous evidence, researchers have done specific studies of the talking head role in learning environments. Homer et al. made an experiment in which participants viewed one of two versions of a computer-based multimedia presentation: One included a lecturer with synchronized slides, and the other his audio narration. The two conditions were compared measuring recall, cognitive load and social presence. A significant difference was found only for cognitive load, with the first condition experiencing greater [5]. Later Kizilcec et al. conducted a similar experiment using eye-tracker to measure information

retention, visual attention, and affect making similar media comparisons. Participants strongly preferred instruction with the face and perceived it as more educational. They dwelled for longer periods on the face than on the slide content, even the area of the face was smaller [1].

The present study proposes an additional view to study the talking head use, considering the changes on mental functioning produced under high levels of cognitive load. It has been found it creates a weaker processing of emotion and attention [6]. Regarding this concern the following hypothesis is tested: The presence or absence of a talking head in an academic video, and its moment of appearance in the learning process affects the emotional and other cognitive states of a user and therefore learning.

III. MATERIALS AND METHODS

A. Experiments design and Materials

The participants of the study were graduate students ($N=12$; between 25 and 33 years old), who were exposed to 3 videos about different types of open software licenses: Aferro, LGPL and BSD. Each participant was exposed to 3 videos, each with a different condition: instructor always presented (talking head condition) (Figure 1), only instructor's voice present (only audio condition), and instructor presented only at the beginning of the video (Mixed condition). Videos were randomized with respect to presentation order, condition, and content.

Figure 1. Screenshot of video with talking head.

The content of the video was created following the instructions provided by an expert in the field and divided into “Introduction”, “Explanations” and “Conclusion” sections, corresponding to different levels of difficulty manipulating the number of items to be remembered and the structure of the information. Additionally, the content was adjusted according to the Flesch Kincaid readability test scores.

The texts in the videos were stylistically similar, built using the same grammatical structures, and paragraphs lengths [2]. They were revised by the expert and later by an experienced educator. The rest of the slides were created using similar configuration of previous experiments [1],

adapting it to the Universitat Pompeu Fabra MOOC's characteristics. Each slide contained no more than 4 sentences with maximum words each. The texts in the videos appeared before the talking head started to speak in order to avoid increasing the attention division [7] through the “pop up effect” produced by the text appearance.

The electroencephalogram (EEG) signal of participants was recorded during the sessions using the EMOTIV EPOC EEG system, that consists of 14 data-collecting electrodes, labeled according to the international 10-20 system, and a wireless amplifier [8].

B. Procedure and measurement

The participants sat in front of a computer showing a presentation containing the instructions and learning material. While subjects read the initial instructions, the researcher placed the electrodes of the EEG on their heads. Participants were asked to read introductory concepts in order to balance the previous knowledge among participants. The EEG signal was recorded only when they watched the videos. After the completion of each video, a recall test with 6 multiple choice questions, and a self-reported assessment of mental effort with the Bratfish, Borg, and Dornic scale [2] was performed. At the end of the experiment, subjects completed a questionnaire to evaluate the social presence level experienced with each media, and made an overall comparison between the conditions assessing emotional states, and learning perception.

For decoding the participants' emotional states from their brain activity, we followed the approach proposed by Ramirez et al. [8], which maps the recorded EEG signal to the arousal-valence plane. Arousal and valence levels were calculated extracting the EEG signal at four locations in the prefrontal cortex: AF3, AF4, F3 and F4.

Arousal level was estimated computing the ratio of the alpha and beta waves at the above 4 locations. Concretely, arousal level is calculated with the following formula:

$$\text{Arousal} = \beta / \alpha. \quad (1)$$

Estimation of valance was based on the fact that left frontal inactivation is an indicator of a withdrawal response, often linked to a negative emotion, and right frontal inactivation may be associated to an approach response, or positive emotion. Concretely, valence values were computed using the following equation: :

$$\text{Valence} = \alpha F4 / \beta F4 - \alpha F3 / \beta F3. \quad (2)$$

For estimating the participant's cognitive load we used “Event-related (de-)synchronization”, that estimates online measures of brain activity. It calculates the perceptual decrease (event-related desynchronization; ERD) or increase (event-related synchronization; ERS) in the alpha band power during the task (activation) interval compared with a

baseline interval [2][3]. The ERD/ERS index is calculated using the following formula:

$$\text{ERD/ERS \%} = 100 * (\text{BaselineIntervalBandPower} - \text{TestIntervalBandPower}) / \text{BaselineIntervalBandPower}. \quad (3)$$

The data was processed in the OpenVibe software and analysed in the software SPSS.

IV. RESULTS

The changes in the computed cognitive load experienced by the subjects within the conditions were tested with the Wilcoxon signed rank test for related samples. There were found statistical significant differences only between the talking head video ($Mdn = 0.65$), and the mixed type ($Mdn = 0.74$, $z = -1.972$, $p < 0.05$).

For analyzing EEG emotional states, it was performed a MANOVA test. No significant differences were found in arousal but there were significant differences found between valence levels: $F(5.169)$ $p < 0.001$ (partial Eta square = 0.070). Using a Bonferroni post-hoc test ($p < 0.001$) it was found that valence was statistically different between the audio narration and the other conditions. Valence was more positive in the audio condition.

Participants' recall, social presence, or opinions comparing the emotional or functional benefits of the conditions did not show any statistical significance difference.

V. CONCLUSIONS

It can be suggested that the increment in cognitive load in the mixed condition was generated by the need learners had to look for a new media for attending the presentation, given the talking head disappearance. This fact could result very demanding due the pre-experienced state of high cognitive load created by the difficulty of the material, which could exhaust mental resources. Consequently it can be concluded that it is recommendable that designers of academic videos make a distinction between high and low content difficulty and cognitive load, and do not make the talking head to appear and disappear when this measurement is expected to be high.

It is interesting to mention that the talking head condition does not provide important learning outcomes compared to the use of audio only. Although there was

found more positive valence states in the audio condition, it is difficult to extract precise conclusions from this result. Additional research with other content and presenters is needed to further study this aspect, contrasting results with related research showing more positive effects when the talking head was present [1]. Future research should also assess the appearance and disappearance of the talking heads in low cognitive load scenarios to provide insides about how to interact with this resource to inform design methods for academic videos.

ACKNOWLEDGMENT

This research has been partially funded by the Spanish Ministry of Economy and Competitiveness, EEE (TIN2011-28308-C03-03), TIMUL (TIN2013-48152-C2-2-R) and RESET (TIN2014-53199-C3-3-R) projects. D.H.L. is a Serra Hnter Fellow. The authors would like to thank the professor Joan Codina for his participation in the content design and recording of the videos.

VI. REFERENCES

- [1] R. Kizilcec, K. Papadopoulos, and L. Sritanyaratana, "Showing Face in Video Instruction: Effects on Information Retention, Visual Attention, and Affect," *Proc. Conference on Human Factors in Computing Systems (CHI 14)*, April 2014, pp. 2095-2102.
- [2] P. D. Antonenko and D. S. Niederhauser, "The influence of leads on cognitive load and learning in a hypertext environment," *Comput. Human Behav.*, vol. 26, no. 2, Mar. 2010, pp. 140-150.
- [3] P. Antonenko, F. Paas, R. Grabner, and T. Gog, "Using Electroencephalography to Measure Cognitive Load," *Educ. Psychol. Rev.*, vol. 22, no. 4, Apr. 2010, pp. 425-438.
- [4] R. E. Mayer and R. Moreno, "Nine Ways to Reduce Cognitive Load in Multimedia Learning," *Educational Psychol.*, vol. 38, no. 1, 2001, pp. 43-52.
- [5] B. D. Homer, J. L. Plass, and L. Blake, "The effects of video on cognitive load and social presence in multimedia-learning," *Computes in Human Behavior*, vol. 24, no. 3, May 2008, pp. 786-797.
- [6] N. Berggren, A. Richards, J. Taylor, and N. Derakshan, "Affective attention under cognitive load: reduced emotional biases but emergent anxiety-related costs to inhibitory control," *Front. Hum. Neurosci.*, May 2013, doi: 10.3389/fnhum.2013.00188
- [7] M. Florax and R. Ploetzner, "What contributes to the split-attention effect? The role of text segmentation, picture labelling, and spatial proximity," *Learn. Instr.*, vol. 20, no. 3, pp. 216-224, Jun. 2010.
- [8] R. Ramirez and Z. Vamvakousis, "Detecting Emotion from EEG Signals Using the Emotive Epoc Device," *Proc. International Conference on Brain Informatics*, Aug. 2012, pp. 175-184.