ScaffoMapping: Assisting Concept Mapping for Video Learners

Shan Zhang[1](✉), Xiaojun Meng[2], Can Liu[1,3], Shengdong Zhao[1], Vibhor Sehgal[1], and Morten Fjeld[4]

[1] NUS-HCI Lab, National University of Singapore, Singapore 117418, Singapore
shan_zhang@u.nus.edu, zhaosd@comp.nus.edu.sg, sehgalvibhor@gmail.com
[2] Noah's Ark Lab, Huawei Technologies, Shenzhen, China
mengxiaojun2@huawei.com
[3] School of Creative Media, City University of Hong Kong, Kowloon, Hong Kong
canliu@cityu.edu.hk
[4] t2i Lab, Chalmers University of Technology, Gothenburg, Sweden
fjeld@chalmers.se

Abstract. Previous research has shown that having learners construct concept maps can bring better learning outcome. However, in video learning scenario, there is not sufficient support for learners to create concept maps from educational videos. Through a preliminary study, we identified two main difficulties video learners face in creating concept maps: navigation difficulty and learning difficulty. To help users to overcome such difficulties, we design scaffolds to assist learners in concept mapping. We present ScaffoMapping, a system aiming for scaffolded concept map creation on educational videos through automatic concept extraction and timestamp generation. Our study, which compares ScaffoMapping with the baseline approach, shows that (1) Learners can create higher quality concept maps with ScaffoMapping. (2) ScaffoMapping enables better learning outcomes in video learning scenario.

Keywords: Concept map · Scaffolding · Video learning

1 Introduction

Concept map is a well-established model of organizing and representing knowledge [27]. It has a graph structure where labeled nodes represent concepts and labeled links represent relationships among concepts. The process of constructing a concept map, which calls "concept mapping", has showed many benefits in facilitating meaningful learning [26]. Compared with rote learning, meaningful learning occurs when students relate new knowledge to their existing knowledge model [7]. Previous research has introduced concept mapping as an effective

Electronic supplementary material The online version of this chapter (https://doi.org/10.1007/978-3-030-29384-0_20) contains supplementary material, which is available to authorized users.

D. Lamas et al. (Eds.): INTERACT 2019, LNCS 11747, pp. 314–328, 2019.
https://doi.org/10.1007/978-3-030-29384-0_20

learning tool in classroom learning environment for decades [35]. Concept maps finished by students can also inform teachers to what extent the learning material is understood by the students, so that they can make adjustments on teaching plans. As such, the exercise of concept mapping can be used as instructional materials and guides and testing vehicles.

As plenty of video-based learning platforms emerge, video has become a popular medium of online education. However, video-based learning has a number of challenges in promoting meaningful learning. For example, novice learners may feel difficult to process new concepts scattered in the video and incorporate them without close guidance from teachers [5]. Moreover, while some videos are well-organized and enjoyable to watch, relying solely on the video content may leave the learners disengaged and lose the chance to reflect on what they have learned [6]. For instructors, gauging learners' experience and comprehension level is also difficult and currently mainly through course reviews, forum posts and quizzes [34].

Previous research has shown that the use of concept map can complement videos in facilitating meaningful learning. Liu et al. [22] showed that editing an existing expert-crafted concept map help learners understand the video content. In their work, the same effects on learning driven by using concept map is observed in video learning context as well, such as reinforcing video learners' understanding by promoting recall, reflection on the video content, and helping learners identify the knowledge gaps between new information and their existing mental model. Finished concept maps also provide a concept-oriented overview of video content, which can further be used as the feedback to video instructors, and future reference to students when preparing for tests.

However, it is not always easy for students to build concept maps from scratch. Novice learners who are unfamiliar with the video content tend to feel frustrated to identify key concepts and their relationships [40]. And the linear representation of video content may bring challenges for learners, which is shown in our preliminary study with 12 participants. We first observed learners building concept maps for video content with a conventional authoring interface (the baseline system in Fig. 1). We saw that they had difficulty in navigating concepts, gaining the overview and video flow, identifying key concepts and relationships.

To mitigate these challenges, we designed scaffolds that aims for easier concept map creation while preserving the learning benefits. We introduce ScaffoMapping, a system to assist the creation of concept maps for video content with automatic concept extraction and automatic timestamp generation. Individual concepts are automatically linked to video timestamps to facilitate content navigation. To evaluate the effectiveness of ScaffoMapping, we conducted a controlled study, comparing ScaffoMapping with the baseline system. Results showed that ScaffoMapping provides better learning outcomes and improves the quality of user-created concept maps. Such system can make concept maps widely accessible and used for online videos.

Overall, the contributions of this paper are:

1. We identified and categorized video learners' difficulties in concept mapping.

2. We designed ScaffoMapping, a system which provides scaffolding help in concept mapping.
3. We empirically evaluated ScaffoMapping by comparing it with the baseline system, and our results showed that ScaffoMapping enables better learning outcomes and higher-quality concept map creation.

2 Related Work

As our work introduces scaffolding concept mapping to facilitate video learning, we review related work on improving the video learning experience, applying concept map in video learning and scaffolding concept mapping.

2.1 Research in Improving Video Learning Experience

Many factors impede video learning platforms in supporting deep and meaningful learning [16] like the lack of personalized instructions and usability problems. Linear representation is one of the biggest problems that prevents learners from exploring video content effectively. Current research provides new ways to improve video learning experience. First, a large body of work focus on the video preview or navigation technique. Some research focuses on creating summaries of segmented video content [29,32,39]. Some research uses visual representation, for example Fisheye [15], video trees [19] and customized word-clouds [38] as navigation aid to help users preview and select some portion of video. However, they either do not provide a concept-oriented view or help learners understand relationships of concepts. NoteVideo [25] identifies conceptual objects to create a summarized image, but is limited to the blackboard-style video. Other research focuses on providing a more engaging learning experience. For example, in-video prompting questions [34] are used to increase learning and help instructors get feedback.

2.2 Applying Concept Map in Video Learning

Previous research has shown the value of providing concept maps in video learning. Some research utilizes the non-linear property of concept map to organize learning materials across different media (e.g., using hierarchical concept maps to support dynamic non-linear learning plans [31]). For integrating concept map within the video content, the most relevant work is done by Liu et al. [22], which utilizes experts or crowd workers to generate interactive concept maps for lecture videos, and has shown the learning benefits on supporting learners' understanding and reflection.

However, previous works focus on the concept maps instead of concept mapping process. They provide complete concept maps generated by crowd workers, experts [22] or Natural Language Processing techniques [40] to learners. While providing complete concept maps saves time, it may put learners in a position where they only passively take in knowledge with little autonomous learning [10], and finally undermine learners' learning performance [17,24]. What's more, editing concept maps by learners are poorly supported in video learning scenario. Extending previous work, we focus on the concept mapping process to facilitate video learning.

2.3 Scaffolds in Concept Mapping

Showing a complete concept map done by crowd workers or experts may result in little autonomous learning. On the contrary, concept mapping by learners themselves can promote autonomous learning and foster deeper processing, but it is time-consuming. What's more, this process requires effort, which usually leads to cognitive overload and finally decreases the learning outcomes as well as learners' motivation [11,20]. Therefore, researchers have been exploring using scaffolds to strike the balance between automation and manual composition [11]. Scaffolds are supports designed to assist learners in mindfully participating in work that would otherwise be too difficult or complex for a novice [13]. Numerous works have explored additional tool-based scaffolds to reduce the cognitive workload in concept mapping [28].

Luchini et al. [23] recognized users' difficulties in concept mapping on handheld computers and designed Pocket PiCoMap to help learners create better concept maps on handheld computers. Cheung et al. [12] proposed a collaborative concept mapping tool, which helps participants co-construct concept maps and improves understanding of the subject matter. Although these tools were successful in some ways, none of them focused on video learning scenario, which probably brings an extra burden for concept mapping. In this research, we focus on the video learning scenario and conduct the following studies.

3 Preliminary Study

To understand how and where learners need to be supported when creating concept maps for videos, we implemented a baseline system (Fig. 1) consisting of a web-based video player for playing lecture video and a canvas for creating concept maps. Learners can watch, pause, replay and navigate the video content. They can use keyboard/mouse shortcuts to create a blank concept, link and label them on the canvas. They can also zoom in/out the canvas according to their needs.

3.1 Study Design

We recruited 12 participants [P1–P12] (5 female) ranging from 22–30 years old from a local university. They were instructed to watch a video (Consciousness-Psychology[1]) and author the concept map in the best quality as they can according to what they had learned. There were no constraints on completion time,

[1] https://www.youtube.com/watch?v=jReX7qKU2yc.

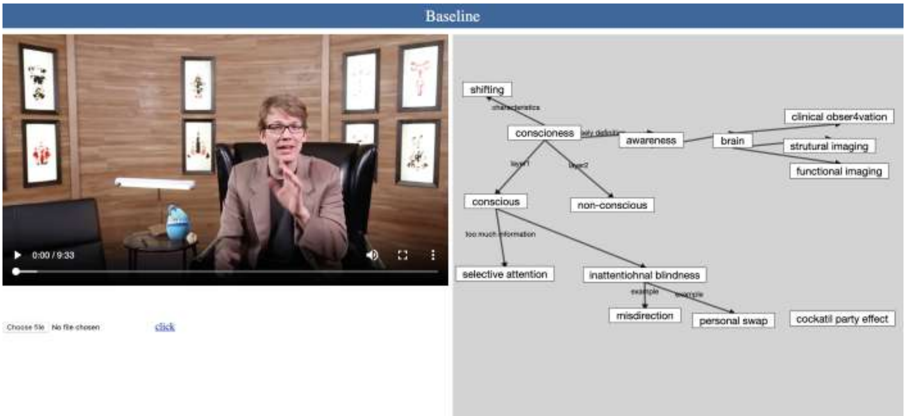


Fig. 1. An example concept map created by P5. The rectangle nodes represent concepts and the directed edges represent the relations. The baseline system interface consisting of a video player and a canvas. Users can create concepts, links and link labels on the canvas.

and they were free to navigate the video (fast-forward or rewind) and author the concept map as they want.

After they finished this task, they were asked to fill in a post-questionnaire. The questionnaire included questions to understand their prior knowledge about the video topic, their engagement level, and perceived difficulty of the video. We also carried out a semi-structured interview for each participant to know more about his/her experience in the experiment.

3.2 Results

Participants reported relatively low levels of prior knowledge on the video topic (10-scale Likert question: 1: never heard about it, 10: understand very well) with M = 4.16 (SD = 2.22). Their engagement while watching the video was high (10-scale Likert question: 1: not engaged at all, 10: Very engaged) with M = 7.5 (SD = 1.76). The perceived difficulty of video (10-scale Likert question: 1: very easy, 10: very difficult) was high with M = 7 (SD = 2.97).

Overall, participants reported that authoring concept maps facilitated their learning of the video content. They were able to remember more concepts compared with only watching the video [P2, P7], understand the relationships among concepts [P2, P9] and understand the video structure [P7]. Some reported that authoring concept maps was similar to taking notes during lectures [P6, P10]. In terms of authoring strategy, most participants noted down concepts during the first time watching as they had limited memory in remembering concepts [P2, P3, P5, P7, P9]. Some watched most of the video to get an overview of the video [P11, P12] and authored concept maps during the second time watching; they regarded authoring concept map as a distraction while watching the video [P8].

All participants reviewed the concept map to add more links and organize the structure. On average, they spent 24 min 46 s (SD = 5 min 16 s) on finishing the task (watching video plus authoring).

Participants reported several difficulties in concept mapping from video content. Based on their experiences, we identified two types of difficulties in concept mapping from video. The first type was the navigation difficulty caused by the linear representation of video content. Gaining an overview and identifying the main topic at the beginning was difficult [P3, P11, P12]. They also struggled to continuously pause the video to type and add concepts. As more concepts were created, they lost the order of the concepts' creation time [P5, P12]. Furthermore, a lot of time was spent on navigating back and forth in the video to find a specific concept or identify relationships between two concepts.

The second type was the learning difficulty in understanding the concepts and their relationships during the concept mapping process. Some learners reported less confidence in finding out important concepts and thus noted down everything they did not understand [P1, P4]. Most learners noted down some important concepts in their mind based on cues, like whether a concept appeared many times [P1, P2, P3, P5] or the video lecturer gave an explanation [P7, P9]. All participants remarked that finding the relationships between concepts was the most difficult part as they could not form the spatial structure and relationships during a single time viewing.

4 System Design

4.1 Designing Scaffolds

Based on the insights gained from the preliminary study, we designed scaffolding solutions to alleviate the navigation and learning difficulty. We searched for the previous literature on scaffolding concept mapping and improving video learning experience. Wang et al. [37] demonstrated the benefits of a hyperlink feature in concept mapping. When the students create concept nodes, the nodes in the concept map are hyperlinked to segments of text that the nodes represent. The hyperlink feature creates a consistent connection between the concept map and the content and enables easier concept map reviewing and editing. In the video learning scenario, previous research also established the effectiveness of linking video timestamps with conceptual objects to facilitate navigation [22,25]. For example, Notevideo [25] identified conceptual objects and linked them with timestamps, which serves as an effective navigation aid. Learners can directly jump to the video frame where that object first appeared. Thus, to mitigate the difficulty learners face in navigating concepts during concept mapping, we proposed the first scaffolding solution as linking concepts that learners created with video timestamps automatically.

Second, many research works demonstrated the benefits of providing students with incomplete concept map templates prepared by experts [10,11]. It creates a quick summary of the learning material and helps in students' understanding of content. In the video learning scenario, audio transcripts [2,4] or visual information [19,38] are used to provide a preview of video content. For example, Yadav et al. [38] provided customized word-clouds, which is an aggregated view of concepts in the video and helps learners understand video content. Thus, to mitigate the difficulty learners face in gaining an overview, we proposed the second scaffolding solution as to extract concepts in the video subtitle to provide a visual preview, which also serves as an incomplete template for concept mapping.

Our other design solutions are based on work done by Liu et al. [22] and Luchini et al. [23]. The concept node size can be bound with the importance of that concept. Relationship hints and time information are also added accordingly. To make them suitable for video scenario, we took an iterative design process by conducting informal pilot studies to get feedback from users. The final design solutions are listed in Tables 1 and 2.

Table 1. Navigation difficulties and scaffolding solutions

Navigation difficulties	Scaffolding solutions
Difficult to navigate concepts	Link concepts with video timestamps
Difficult to gain an overview	Extract concepts to provide a visual preview
Difficult to understand the video flow	Bind concepts' color with time information

Table 2. Learning difficulties and scaffolding solutions

Cognitive difficulties	Scaffolding solutions
Difficult to identify important concepts	Bind concepts' size with frequency
Difficult to find the relationship	Provide the relationship hints

4.2 System Implementation

To support scaffolds design, we utilized video and its subtitle file as input. The subtitle file is a text file in SRT or VTT format [18]. A single subtitle block, also called a cue, consists of three parts: the cue index, the timing information and the text (Fig. 2). The text feature of the subtitle has been widely used in identifying video topics to support video classification [8], and the timing information is usually combined with visual information to improve the interactivity of video [18,33].

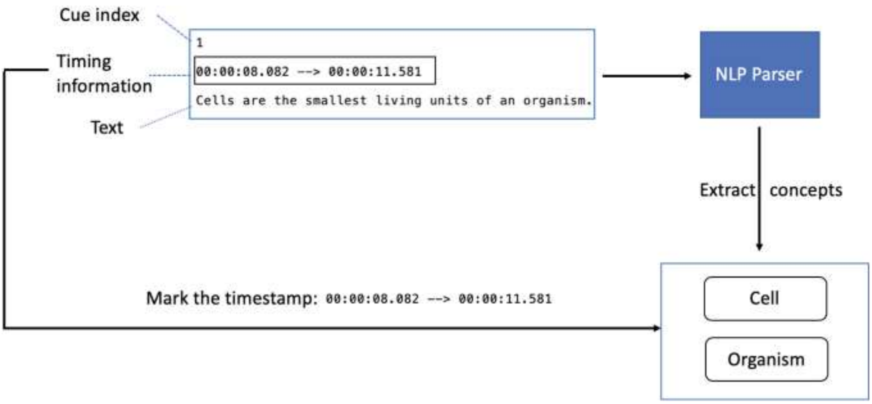


Fig. 2. An example of a subtitle block (the cue index, the timing information and the text) and ScaffoMapping work flow

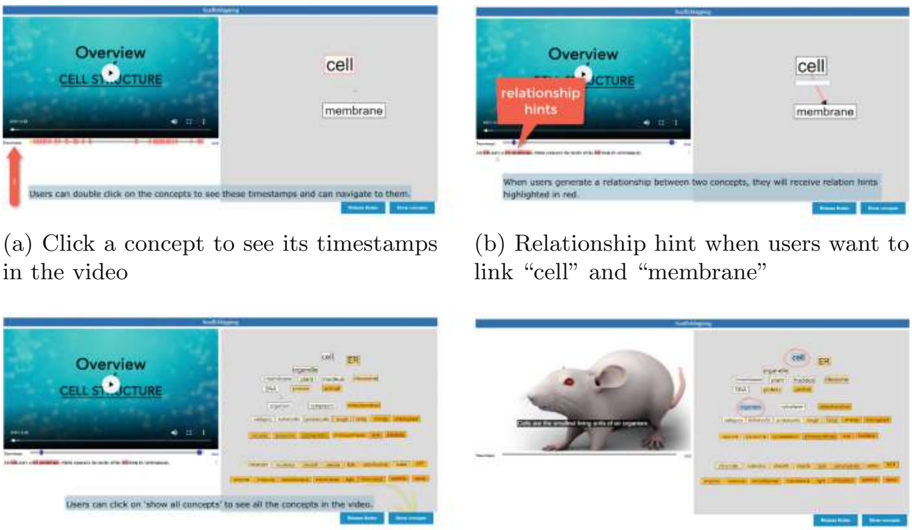
Automatic Concept Extraction: We first used Google Natural Language API [1] to do text analysis and extract potential key concepts from subtitles. This API was used to analyze the structure of sentences and identify proper nouns in subjects or objects. These nouns are often regarded as key concepts. The Great Noun List [3] was used to identify the stems that are related to a given topic. Those extracted nouns (or stems) are more likely to contain the meaning that relates to the topic in a video. After generation of the concepts, we calculate the frequency that each concept appears in the subtitle. The more frequent concepts get larger font sizes and rectangle sizes.

Automatic Timestamp Generation: Each concept was coded with the time information of the subtitle cue. As a concept may appear more than once in the subtitle, after we marked all the timestamps of the concept, we marked the first one as the time of creation, which was the yellow dot in Fig. 3(a). Each concept was automatically assigned with a color from white to orange according to their creation time. Learners could click a concept to see all the moments it appeared in the video (Fig. 3(a)) and double click to navigate to the creation time. We also visualized in-video progress by highlighting mentioned concepts in every frame (blue colored concepts in Fig. 3(d)).

Concept relationships were made explicit by the hierarchical structure of the concept map as follows: We first sorted the concepts according to their frequency and creation time (Fig. 3(c)), then developed relationship hints, so that when a user linked two concepts in the same sentence, the text of that sentence would show below the video and the timestamp of the sentence would be marked in the timeline (Fig. 3(b)). The system would give the hint "No direct link" if none occurred, which means the two concepts users want to link don't appear in the same sentence. In this case, they need to build the relationship based on their own understanding. Keyboard shortcuts and zoom in/out were provided in the same way as in the baseline system. A release nodes function allowed users to remove all isolated nodes and keep linked ones.

5 Evaluation Study

ScaffoMapping is designed to facilitate video learning through scaffolding concept mapping for video learners, so we are interested in knowing: (1) Will the learners be able to create better concept maps with the help of ScaffoMapping interface? (2) How's the learning outcome in ScaffoMapping compared with the baseline interface? With these questions in mind, we conducted a user study to compare ScaffoMapping with the baseline system.



(a) Click a concept to see its timestamps in the video

(b) Relationship hint when users want to link "cell" and "membrane"

(c) Display all the concepts for reference

(d) Mentioned concepts are highlighted in blue as video plays

Fig. 3. The ScaffoMapping interface integrates interactive concept mapping into the video player. Each concept users created is automatically linked to its timestamps in the video. Relationship hints are given when users create links. All the concepts are displayed as the reference and lecture progress is visualized in blue. (Color figure online)

5.1 Study Design

Participants and Materials: We recruited 12 participants [P1–P12] (6 male and 6 female) ranging from 21–27 years old from a local university. We selected two videos on two different topics from the same YouTube education channel (V1: Immune System[2], V2: Consciousness[3]) with similar styles and similar lengths

[2] https://www.youtube.com/watch?v=GIJK3dwCWCw.
[3] https://www.youtube.com/watch?v=jReX7qKU2yc.

(V1: 9:12; V2: 9:32). Participants rated their prior knowledge level through a 10-point Likert Scale (1: Never heard about it, 10: understand very well) and perceived difficulty of video content (1: very easy, 10: very difficult) after watching them. As the participants were neither from psychology nor from biology major, there was little risk that they were familiar with the advanced content of the videos. This was consistent with their self-reported similar level of prior knowledge (V1: $M = 3$, $SD = 2.37$; V2: $M = 3.25$, $SD = 1.54$ ($t(11) = 0.491$, $p = 0.633$)). Videos were pilot tested with three users to ensure the similar difficulty level, which was consistent with participants' perceived difficulty level (V1: $M = 6.08$, $SD = 2.06$; V2: $M = 5.08$, $SD = 2.31$ ($t(11) = 0.964$, $p = 0.356$)).

Tasks and Procedure: We began by providing introductions of the two systems to each participant: baseline version and scaffolds version. Before starting the experiment, participants had time to practice and get familiar with the two systems. Then each participant experienced two experimental sessions where they watched the videos (V1, V2) and created concept maps using both systems (baseline, ScaffoMapping). The order of the systems was counter-balanced while the order of the video remained the same (e.g., P1 watched V1 using baseline system, then watched V2 using ScaffoMapping; P2 watched V1 using ScaffoMapping, then watched V2 using baseline system). They had the freedom to author concept maps during or after watching the videos, pause and resume, and watch multiple times, but the time for each session was limited to be the video duration plus 10 min. They could not take further notes such as with paper and pencil.

5.2 Analysis

Concept Map Quality Evaluation: To measure the quality of the concept maps, two domain experts were invited for each video to evaluate the concept maps generated from the baseline system and from ScaffoMapping. They watched the videos and evaluate the validity of concepts and relationships (links and link phrases). Adopted from [21], one valid concept or relationship was given 1 point. Invalid concepts and relationships were those containing typos or were incorrect, as determined by the experts. The interrater reliability correlations of the experts for V1 and V2 were 0.94 and 0.96, which indicates high reliability.

Learning Outcome Evaluation: To measure the learning gain, a 5-question test was administered to each participant after each session. The questions consisted of 3 True or False questions and 2 multiple choice questions from the video content. For example, for V1 on immune system, Q1: "The skin only supplies simple physical protection. T/F" Q4: "Which one of the following is not the internal defense? A. Phagocytes B. Mucous membranes C. Antimicrobial proteins D. Neutrophils". Participants did not have access to videos and concept maps during the test. After they had finished the two experiment sessions, we conducted a semi-structured interview asking their system preference and experiences.

Participants were told in advance that their performance indicator was the total of test scores (5 points) plus the concept map scores (convert concept map quality to a full-score of 5 points). They received $10 for finishing the experiment and an extra $10 for the best performance.

6 Results

Although the number of participants is relatively small, effects in the valid concept/link number, learning outcome and user preference are observed with statistical significance.

6.1 Concept Map Quality

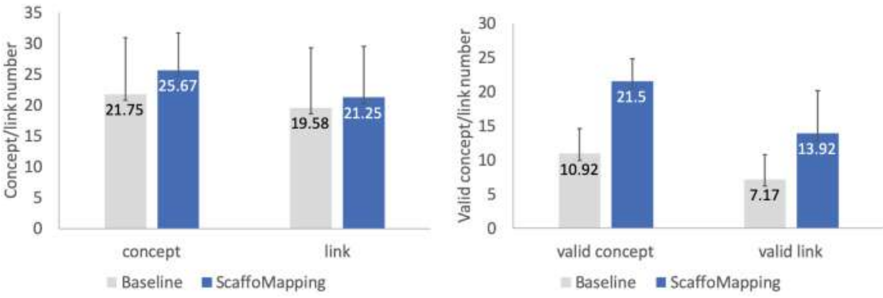
We calculated the numbers of concepts and links participants generated. Two domain experts for each video rated the number of valid concepts and links as we described earlier. The summary of results are shown in Fig. 4(a) and (b).

There was no significant difference in the total numbers of concepts and links in the baseline system and ScaffoMapping. However, there was significant difference in the numbers of valid concepts ($t(11) = 8.072$, $p < 0.001$) and valid links ($t(11) = 3.906$, $p < 0.01$). This result shows that ScaffoMapping can facilitate learners to create concept maps with higher quality concepts and links with the scaffolding solutions provided.

6.2 Learning Outcome Test

Overall, participants using ScaffoMapping ($M = 3.92$, $SD = 1.38$) outperformed baseline systems ($M = 3$, $SD = 1.21$) in the tests. A paired t-test revealed significant improvement in the test score ($t(11) = 3.19$, $p < 0.01$). Our results showed that ScaffoMapping assists effective video learning.

There are two reasons explaining the learning benefits of ScaffoMapping. First, ScaffoMapping supports efficient navigation through the automatic timestamp generation. Considerable time was saved and users could focus on construction work which involves more analytical thinking, e.g. identifying relationships. According to the interview result, all participants reported that they like the automatic timestamp generation, which provides the hyperlink from concepts to the video content. Second, with the extracted concepts, users reported a more positive learning experience when they found the gap between existing concepts and their mental model. With the template concepts provided, participants try to link between the template concepts and the knowledge they learned in the video, e.g. "I pay attention to the concepts and try to make sense of them" [P11], "I can identify the missing gaps and digest the information on the spot." This was consistent with the learning theories [9], which shows that meaningful learning is facilitated when learners integrate new concepts and propositions into their existing cognitive structure.



(a) (b)

Fig. 4. (a) Mean of concept and link number of baseline system and ScaffoMapping. The result shows no significant difference. (b) Mean of valid concept and link number of baseline system and ScaffoMapping. The result shows learners use ScaffoMapping can create more valid concepts and links.

6.3 User Preference

We collected participants' feedback through the semi-structured interview. The majority of participants preferred ScaffoMapping (10/12) over the baseline system. Participants reported that ScaffoMapping made the authoring process easier and more efficient. The extracted concepts remove the burden of tedious typing and makes them focus on constructing links between concepts. The highlighting feature visualizes the in-video progress and can help them find connection easier [P3, P5, P7, P12]. And the automatic timestamp linking helps them navigate and review concept maps easily. They also indicated that they would like to save the concept maps as interactive notes for the video, and refer to them when they want to review or prepare for exams.

The remaining two participants [P4, P10] reported that they wanted a combination of baseline system plus automatic timestamp generation feature. In ScaffoMapping, they were confused by the extracted concepts appearing in the beginning, and after they identified key concepts, they needed to check their existence in the canvas, which created extra burden for them. While the baseline system gave them a sense of more "control", navigating back and forth to find the concept was tedious. Thus, they preferred to have a combination system.

7 Limitation and Future Work

We conducted an interview to gather feedback from learners about their video learning experience with ScaffoMapping. While the majority gave positive feedback to ScaffoMapping, two participants reported an expectation of baseline system combining automatic timestamp generation. It is still unclear whether expertise or domain knowledge has an influence on the choice of scaffolding systems. Due to the small size of participants, no correlation between system preference and prior knowledge with test performance was found. Furthermore, since our participants were recruited from the university community, more advanced scaffolding may be needed for less skilled learners. While some design solutions may solve this problem to some extent (e.g. triggering the automatic concept extraction by user control), it is worth investigating designing personalized scaffolding strategies according to learners' knowledge level to maximize video learning results. For example, combining with psychophysiological computing to measure the learners' frustration or attention to provide scaffolds at proper time [14].

We utilized video subtitle as the text source to provide relationship hints, which is supported by previous research that speech is the main carrier of information in nearly all lectures [39]. Inspired by the TF-IDF indices [30], if two concepts appeared in the same sentence(s), the sentence(s) would be displayed as the hint to help learners identify relationships. However, this feature can be interfered by pronouns like "it", or "them". For example, if the learner want to link concept A and B whose relationship is shown in sentence C, but A is referred as "it", the system will fail to identify sentence C as the relationship hint. To provide high-quality relationship hints, external sources or databases like Wikipedia [36] can be added to help with the relationship analysis.

While learner-generated concept maps can provide valuable feedback for video lecturers, analyzing concept maps can bring an extra burden for lecturers especially when a video has a large number of learners. To make ScaffoMapping more useful for lecturers and educational video developers, automatic concept map evaluation could be explored as an extension work. Finally, our evaluation study focused on video learners. In the future, we will extend this work to larger group of people. We expect to see how teachers, practitioners, educational video developers make use of this system to support learners.

8 Conclusion

In this paper, we proposed scaffolding solutions to assist concept mapping for video learners. We implemented these solutions in the ScaffoMapping system. Our evaluation study indicated that learners can create concept maps with better quality, as well as obtaining higher learning outcomes using ScaffoMapping. Our investigation on scaffolding solution shows promising benefits of semi-automation in learning systems. We encourage future research on similar solutions and believe this will have a wide impact on digital learning communities.

Acknowledgements. This research was funded by National University of Singapore Academic Research Fund T1 251RES1617. We thank Philippa Beckman and Barrie James Sutcliffe for proofreading, and Samangi Wadinambi Arachchi for her generous help with designing Fig. 4.

References

1. Cloud natural language API. https://cloud.google.com/natural-language/docs
2. Edx. https://www.edx.org
3. The great noun list. http://www.desiquintans.com/nounlist
4. Khan academy. https://www.khanacademy.org
5. Ally, M.: Foundations of educational theory for online learning. Theory Pract. Online Learn. **2**, 15–44 (2004)
6. Anderson, T., Kanuka, H.: Online social interchange, discord, and knowledge construction. J. Distance Educ. **13**, 57–74 (1998)
7. Ausubel, D.P.: The psychology of meaningful verbal learning (1963)
8. Brezeale, D., Cook, D.J.: Automatic video classification: a survey of the literature. IEEE Trans. Syst. Man Cybern. Part C (Appl. Rev.) **38**(3), 416–430 (2008)
9. Cañas, A.J., Novak, J.D.: Concept mapping using Cmaptools to enhance meaningful learning. In: Okada, A., Shum, S.B., Sherborne, T. (eds.) Knowledge Cartography. AIKP, pp. 25–46. Springer, London (2008). https://doi.org/10.1007/978-1-84800-149-7_2
10. Chang, K.E., Sung, Y.T., Chen, I.D.: The effect of concept mapping to enhance text comprehension and summarization. J. Exp. Educ. **71**(1), 5–23 (2002)
11. Chang, K.E., Sung, Y.T., Chen, S.F.: Learning through computer-based concept mapping with scaffolding aid. J. Comput. Assist. Learn. **17**(1), 21–33 (2001)
12. Cheung, L.S.: A constructivist approach to designing computer supported concept-mapping environment. Int. J. Instr. Media **33**(2), 153–165 (2006)
13. Collins, A., Brown, J.S., Newman, S.E.: Cognitive apprenticeship: teaching the crafts of reading, writing, and mathematics. In: Knowing, Learning, and Instruction: Essays in Honor of Robert Glaser, vol. 18, pp. 32–42 (1989)
14. Dirican, A.C., Göktürk, M.: Psychophysiological measures of human cognitive states applied in human computer interaction. Procedia Comput. Sci. **3**, 1361–1367 (2011)
15. Divakaran, A., Forlines, C., Lanning, T., Shipman, S., Wittenburg, K.: Augmenting fast-forward and rewind for personal digital video recorders. In: Consumer Electronics, pp. 43–44 (2005)
16. Garrison, D.R., Cleveland-Innes, M.: Facilitating cognitive presence in online learning: interaction is not enough. Am. J. Distance Educ. **19**(3), 133–148 (2005)
17. Griffin, C.C., Malone, L.D., Kameenui, E.J.: Effects of graphic organizer instruction on fifth-grade students. J. Educ. Res. **89**(2), 98–107 (1995)
18. Hu, Y., Kautz, J., Yu, Y., Wang, W.: Speaker-following video subtitles. ACM Trans. Multimedia Comput. Commun. Appl. (TOMM) **11**(2), 32 (2015)
19. Jansen, M., Heeren, W., van Dijk, B.: Videotrees: improving video surrogate presentation using hierarchy. In: International Workshop on Content-Based Multimedia Indexing, CBMI 2008, pp. 560–567. IEEE (2008)
20. Katayama, A.D., Robinson, D.H.: Getting students partially involved in note-taking using graphic organizers. J. Exp. Educ. **68**(2), 119–133 (2000)
21. Kwon, S.Y., Cifuentes, L.: The comparative effect of individually-constructed vs. collaboratively-constructed computer-based concept maps. Comput. Educ. **52**(2), 365–375 (2009)
22. Liu, C., Kim, J., Wang, H.C.: ConceptScape: collaborative concept mapping for video learning. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, p. 387. ACM (2018)

23. Luchini, K., et al.: Scaffolding in the small: designing educational supports for concept mapping on handheld computers. In: CHI 2002 Extended Abstracts on Human Factors in Computing Systems, pp. 792–793. ACM (2002)
24. McCagg, E.C., Dansereau, D.F.: A convergent paradigm for examining knowledge mapping as a learning strategy. J. Educ. Res. **84**(6), 317–324 (1991)
25. Monserrat, T.J.K.P., Zhao, S., McGee, K., Pandey, A.V.: NoteVideo: facilitating navigation of blackboard-style lecture videos. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1139–1148. ACM (2013)
26. Novak, J.D.: Concept maps and Vee diagrams: two metacognitive tools to facilitate meaningful learning. Instr. Sci. **19**(1), 29–52 (1990)
27. Novak, J.D., Cañas, A.J.: The theory underlying concept maps and how to construct and use them (2008)
28. Oliver, K.: An investigation of concept mapping to improve the reading comprehension of science texts. J. Sci. Educ. Technol. **18**(5), 402–414 (2009)
29. Pavel, A., Hartmann, B., Agrawala, M.: Video digests: a browsable, skimmable format for informational lecture videos. In: Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology, pp. 573–582. ACM (2014)
30. Ramos, J., et al.: Using TF-IDF to determine word relevance in document queries. In: Proceedings of the First Instructional Conference on Machine Learning, vol. 242, pp. 133–142 (2003)
31. Schwab, M., et al.: Booc.io: an education system with hierarchical concept maps. IEEE Trans. Vis. Comput. Graph. **23**(1), 571–580 (2017)
32. Seidel, N.: Making web video accessible: interaction design patterns for assistive video learning environments. In: Proceedings of the 20th European Conference on Pattern Languages of Programs, p. 17. ACM (2015)
33. Shin, H.V., Berthouzoz, F., Li, W., Durand, F.: Visual transcripts: lecture notes from blackboard-style lecture videos. ACM Trans. Graph. (TOG) **34**(6), 240 (2015)
34. Shin, H., Ko, E.Y., Williams, J.J., Kim, J.: Understanding the effect of in-video prompting on learners and instructors. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, p. 319. ACM (2018)
35. Taylor, B.M., Beach, R.W.: The effects of text structure instruction on middle-grade students' comprehension and production of expository text. Reading Res. Q. **19**, 134–146 (1984)
36. Wang, F., Li, X., Lei, W., Huang, C., Yin, M., Pong, T.-C.: Constructing learning maps for lecture videos by exploring Wikipedia knowledge. In: Ho, Y.-S., Sang, J., Ro, Y.M., Kim, J., Wu, F. (eds.) PCM 2015. LNCS, vol. 9314, pp. 559–569. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24075-6_54
37. Wang, S., Walker, E., Wylie, R.: What matters in concept mapping? Maps learners create or how they create them. In: André, E., Baker, R., Hu, X., Rodrigo, M.M.T., du Boulay, B. (eds.) AIED 2017. LNCS (LNAI), vol. 10331, pp. 406–417. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-61425-0_34
38. Yadav, K., et al.: Content-driven multi-modal techniques for non-linear video navigation. In: Proceedings of the 20th International Conference on Intelligent User Interfaces, pp. 333–344. ACM (2015)
39. Yang, H., Siebert, M., Luhne, P., Sack, H., Meinel, C.: Automatic lecture video indexing using video OCR technology. In: 2011 IEEE International Symposium on Multimedia (ISM), pp. 111–116. IEEE (2011)
40. Zubrinic, K., Kalpic, D., Milicevic, M.: The automatic creation of concept maps from documents written using morphologically rich languages. Expert Syst. Appl. **39**(16), 12709–12718 (2012)