



A new application of the temporal contiguity effect in designing narrated slideshows

Tzu-Chien Liu¹ · Yi-Chun Lin¹ · Fred Paas^{2,3}

Accepted: 19 December 2021 / Published online: 30 January 2022
© Association for Educational Communications and Technology 2021

Abstract

We investigated whether the temporal contiguity effect, which holds that information sources, such as visual information and narration need to be temporally coordinated for learning to be effective, can also be found in narrated slideshows. A concurrent presentation-key point format (CPK), in which visual information sequentially appeared as key points on the slide with corresponding narration, was compared to a concurrent presentation-whole format (CPW), in which visual information was shown all at once on the slide with corresponding narration, and a sequential presentation format (SP), in which the narration was played first before all the corresponding visual information was presented at once. Ninety-nine undergraduates were randomly divided across the CPK, CPW and SP conditions. Results revealed that participants in the CPK group had higher post-test performance and learning efficiency than participants in the CPW and SP conditions. Performance in the CPW condition was higher than in the SP conditions, but only in terms of learning efficiency. The results suggested that the occurrence of the temporal contiguity effect not only depends on whether the presentation of narration and visual information in narrated slideshows is concurrent or not, but also on how concurrent it is.

Keywords Cognitive load theory · Temporal contiguity effect · Narrated slideshows · Multimedia

Introduction

Digital learning has been a main solution for learning especially during the recent COVID-19 pandemic (Mulenga & Marbán, 2020). As one of the most common multimedia materials, narrated slideshows are widely used in different types of digital

✉ Tzu-Chien Liu
tzuchien@ntnu.edu.tw

¹ Department of Educational Psychology and Counseling, Institute for Research Excellence in Learning Sciences, National Taiwan Normal University, 162, HePing East Road, Section 1, Taipei, Taiwan

² Department of Psychology, Education, and Child Studies, Erasmus University Rotterdam, Rotterdam, The Netherlands

³ School of Education/Early Start, University of Wollongong, Wollongong, Australia

learning environments, such as flipped classrooms (DeLozier & Rhodes, 2017) and Massive Open Online Courses (MOOCs) (Zhuhadar et al., 2015). Narrated slide-shows which present static visual information accompanied by synchronized narration through computer are typically created with programs such as PowerPoint, Keynote or Prezi. A unique characteristic of narrated slide-shows is that verbal information is presented in a dynamic way (i.e., narration), but visual information on the slide is often presented in a static way (i.e., static graphics). It's very different from another often used format for digital learning, narrated animation, in which both verbal and visual information are all presented in a dynamic way (i.e., animation with narration) (e.g., Liu et al., 2021; Mayer & Anderson, 1991; Mayer et al., 1999).

Spatial and temporal integration of verbal and visual information are considered critical factors for successful learning from multimedia materials composed of the two sources of information (Kalyuga, 2009; Mayer, 2017; Sweller et al., 2019). Research has revealed clear guidelines for enhancing visual-verbal information integration in the learning from narrated animation (e.g., Ayres & Sweller, 2005; Mayer & Sims, 1994). Research into the design of narrated slide-shows which visual information presented in a static way, however, has received much less attention.

The design guidelines that have been identified as effective in narrated animation may not be effective in narrated slide-shows, because the two presentation formats differ in visual-information features, which is expected to result in different cognitive processes and levels of cognitive load. On the one hand, when dealing with dynamic visual information in narrated animation, both the visual and verbal information may have disappeared before learners had time to adequately process it and link it to newly appearing information. According to cognitive load theory, this process may create cognitive overload and negatively affect learning (Ayres & Paas, 2007, 2009; Sweller, 2020; Sweller et al., 2011). When dealing with static visual information in narrated slide-show, learners are expected to have more time to link the narration and visual information because the information remains visible on the slide.

On the other hand, narrated slide-shows differ from narrated animations in the level of synchronization of visual information and narration. Narrated animations have a higher level of synchronization because the material with dynamic visual information and narration are usually divided into small separate parts (e.g., Mayer & Anderson, 1991; Mayer et al., 1999). The corresponding specific part of animation is displayed once it is mentioned in the narration. According to this, participants dealing with materials with animation and narration can save cognitive resources in searching and matching the information from two sources. On the contrary, narrated slide-show has a lower level of synchronization because the visualizations are usually shown in a whole picture or table so that the participants need to search the specific part of graphics when they hear the narration. In that way, more cognitive resources need to be spent on visual search-and-match processes in the narrated slide-show, which may result in cognitive overload (Liu et al., 2022).

Because of these differences in cognitive processes and the limited research attention for improving effective integration of verbal and visual information in narrated slide-shows, it is deemed important to investigate whether instructional design principles for narrated animation research are also applicable to narrated slide-shows. Among the possible instructional design principles, the temporal contiguity effect is an interesting one to study (Khacharem et al., 2020; Mutlu-Bayraktar et al., 2019).

Temporal contiguity effect

The temporal contiguity effect, which has been systematically studied in multimedia learning (for a meta-analysis see, Ginns, 2006; see also, Mammarella et al., 2013; Mayer & Anderson, 1991; Mayer et al., 1999; Schüler et al., 2012), holds that better learning performance is achieved when the presentation of verbal and visual information is temporally coordinated instead of temporally separated. Research has indicated that the temporal split-attention may occur if learners need to mentally integrate and process multiple sources of information which are presented sequentially instead of simultaneously (Ayes & Sweller, 2005). Due to the resulting temporal separation in the sequential presentation format, visual search-and-match and cross-referencing between different information sources may substantially increase the burden on working memory and inhibit learning (Sweller et al., 2011). Instead, temporal coordination is the solution for reducing the negative effects from temporal separation (e.g., Mayer & Anderson, 1991, 1992; Mayer & Sims, 1994). For example, Mayer and Anderson (1991) compared the learning performances of a split-attention condition in which the narration was presented before the animation, and an integrated condition in which the narration was presented concurrently with animation. Participants in the integrated condition outperformed participants in the split-attention condition on the test, indicating a temporal contiguity effect. Mayer et al. (1999) further investigated the temporal contiguity effect using different lengths of the narrations and animations. Results showed that the temporal contiguity effect only occurred in animations and narrations with the longer length. In summary, previous research results indicated that temporal contiguity effect is not only related to the concurrent presentation of verbal and corresponding visual information, but also influenced by the length of narrations and animations (e.g., Mayer & Anderson, 1991, 1992; Mayer & Sims, 1994; Mayer et al., 1999).

Theories related to temporal contiguity effect

Cognitive load theory (CLT: Ayres & Sweller, 2005; Sweller et al., 1998, 2019), the cognitive theory of multimedia learning (CTML: Mayer, 2017; Mayer & Anderson, 1991, 1992; Mayer & Sims, 1994), and the related dual-coding theory (DCT; Clark & Paivio, 1991; Mayer & Anderson, 1991) have been important theoretical bases for the design of multimedia learning material, and more specifically for the explanation of the temporal contiguity effect.

According to DCT and the CTML, successful learning from multimedia learning material originates from verbal and visual information referential connections and representational connections. Verbal referential connections are constructed through the cognitive process of verbal coding in which external verbal representations are transformed into internal verbal representations. Visual referential connections are constructed through the cognitive process of visual coding in which external visual representations are transformed into internal visual representations. Representational connections need to be constructed between the verbal and visual representations. Presenting narration with animation can facilitate the connection of the verbal and visual representations (Mayer & Anderson, 1991). In terms of CLT, when the verbal and visual information are presented successively, unnecessary cognitive resources are consumed during holding the verbal information to search for the corresponding elements in the visual information (Owens & Sweller, 2008). Therefore, it is considered better to present two sources of information simultaneously,

which requires less cognitive resources for integrating verbal and visual information, than sequentially.

The present study

In this study, we investigated whether the temporal contiguity effect, which is commonly observed in learning from narrated animations, can also be observed in learning from narrated slideshows. In addition, by comparing different types of narrated slideshows we aimed to identify a way to effectively reduce visual search-and-match load caused by the lower synchronization level of verbal information and visual information in narrated slideshow.

Three popular formats of narrated slideshow were compared to investigate the mentioned research issue. The first format is concurrent presentation-key point (CPK), in which key points of visual information appear one by one on the slide with corresponding narration. The second format is the conventional concurrent presentation-whole (CPW), in which all visual information is presented at once on the slide with corresponding narration. The third format is the sequenced presentation (SP), in which the narration is played first before all the corresponding visual information appears. CPK and CPW are concurrent presentation formats, while SP is a non-concurrent presentation format. In addition, compared with CPW, CPK has a higher synchronization level between visual information and corresponding narration.

Research questions and hypotheses

This study focused on answering the following two research questions:

Q1: Does concurrent presentation in the CPK and CPW conditions lead to higher post-test-scores, lower cognitive load and higher learning efficiency than sequential presentation in the SP condition?

Q2: Does the higher synchronization level of concurrent presentation in the CPK conditions lead to higher post-test-scores, lower cognitive load and higher learning efficiency than the lower synchronization level of concurrent presentation in the CPW condition?

The corresponding hypotheses for the mentioned research questions were:

H1: Participants in the two concurrent presentation conditions (CPK and CPW) would have higher post-test scores, perceive lower cognitive load, and show higher learning efficiency than those in the sequenced presentation condition (SP).

H2: Participants in the CPK condition would have higher post-test scores, perceive lower cognitive load, and show higher learning efficiency than the participants in the CPW condition.

Research method

Experimental design and participants

A single factor experimental design was used to address the mentioned research hypotheses. Participants were 99 undergraduates (24 males, 75 females) with a mean age of 20 years from a university in northern Taiwan. They were randomly assigned to the

Table 1 Participants' characteristics table

Demographic characteristics	CPK		CPW		SP	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Gender						
Male	9	29	7	21	5	16
Female	22	71	26	79	27	84
Major						
Liberal Arts	9	29	8	24	9	28
Social science	16	52	19	58	18	56
Science	6	19	6	18	5	16
Class of year						
First-year	5	16	3	9	6	19
Second year	8	26	10	30	10	31
Junior	14	45	12	36	8	25
Senior	4	13	8	24	8	25

**Fig. 1** Example of two successive slides of the learning material

following three conditions: concurrent presentation-key point (CPK), concurrent presentation-whole presentation (CPW) and sequenced presentation (SP), 33 in each condition. Two participants in the CPK condition and one participant in the SP condition were identified as outliers and eliminated from the data analysis as they deviated more than 2.40 SDs from the mean of the post-test in the assigned condition. The final number of participants in the three conditions was 31 in the CPK condition, 33 in the CPW condition and 32 in the SP condition. Table 1 shows the participants' gender, class of year, and major.

Learning materials

The topic of the learning material was “Mindfulness”. The concepts, related empirical studies, and the brain basis for mindfulness were introduced in Mandarin Chinese through the mentioned three condition-specific versions of narrated slideshow. All versions of narrated slideshow were created with Microsoft PowerPoint. Figure 1 is an example of two successive slides of the learning material. The contents of the narration and visual information for the three versions were identical. In order to present a common design of narrated slideshow, there were twenty-seven slides (7 min 19 s for the CPK and CPW; 9 min 4 s for the SP) in each version of which 10 slides were specifically designed for the conditions.

Both concurrent versions (CPK and CPW) had the same number of narrative words and time duration in the 10 slides (800 words; 3 min 08 s). The sequenced presentation version (SP) contained more narrative words with a longer play duration than the concurrent versions (886 words; 4 min 53 s) because some words were used to connect the narration and the subsequent displayed visual information. So, these words were not related to the content. One example of these words was “the example illustration will be shown in the following slide”.

The three versions differed from each other in terms of the synchronization of the narration and the visual information in the selected slides. Figure 2a–c provides the examples for the three presentation formats in the different versions. For the CPK version, the visually presented key points were shown on the slide one by one with the corresponding narration. The presented key points remained visible on the slide until all key points on the slide were shown. For the CPW version, the visual information within the slide was shown in static format all at once and the corresponding narration was played concurrently. For the SP version, the narration for introducing the visual information was played first and the corresponding visual information within the slide was presented on the slide all at once immediately after the end of the narration.

Measures

A prior knowledge test was carried out to measure participants’ knowledge about the basic concepts of mindfulness and the brain basis for mindfulness before the experiment. The test consisted of 10 multiple-choice questions (e.g., “Which one of the following statements

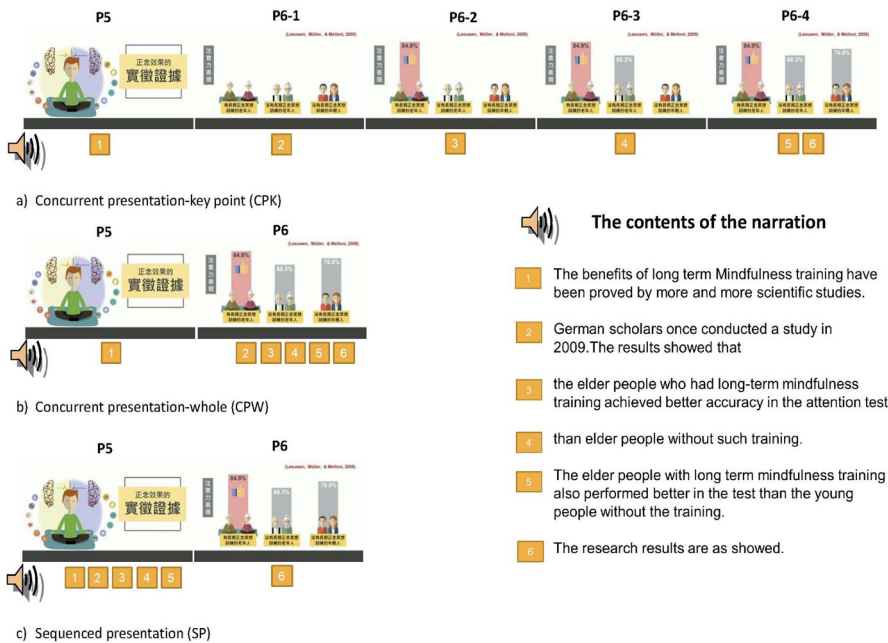


Fig. 2 Examples for the three presentation formats in the different versions (The narration is translated from Mandarin Chinese)

about mindfulness is true?"; "In the structure of the brain, which one of the following brain areas is particularly relevant to emotions?"). 1 point was given for a correct answer and 0 point was given for an incorrect answer. Accordingly, the highest possible score that could be achieved was 10 points.

A *post-test* was used to measure participants' comprehension of the learning materials. The post-test was developed according to the contents from the 10 slides that were specifically designed for the conditions. Two experts on the topic of mindfulness and two test development experts were invited to review the accuracy of all items and the correspondence of each item to learning contents. They were requested to give their comments and suggestions on each item. The research team revised those items according to the comments and suggestions. Some items that needed to be reviewed again by the experts were returned to the experts until the validity of each item was confirmed by all experts. The final version of the test consisted of 31 items including 9 matching items (e.g., "The following picture shows the structure of the inner side of the brain. Please match the brain parts and their names.") and 22 multiple-choice questions (e.g., "In the results of the mindfulness study of Leeuwen et al. (2009), which of the following statement regarding the attention performance of the participants is true?"). The scoring rule is the same as that in the prior knowledge test. 1 point was given for a correct answer and the maximum total points for the post-test is 31, and the lowest possible score was 0. The Kuder-Richardson (KR) 20 (i.e., KR-20) formula was used to test the internal consistency reliability of post-test. The result revealed good consistency reliability for the post-test, KR-20 was .91 (Fraenkel et al., 2011).

A *mental effort rating scale* was used to measure how much mental effort participants invested to comprehend the learning contents of the narrated slideshow. Mental effort is defined as "the aspect of cognitive load that refers to the cognitive capacity that is actually allocated to accommodate the demands imposed by the task" (Paas et al., 2003, p. 63), which is generally acknowledged as an index reflecting the cognitive load (Paas & Van Merriënboer, 1994). A mental effort rating scale which was revised from the scale by Paas and Van Merriënboer (1994) was applied in this study. The Likert type nine-point scale ranged from 1. very, very low mental effort to 9. very, very high mental effort.

Learning efficiency was used to measure the relation between learning time and post-test performance (Liu et al., 2013). Learning efficiency was calculated according to an adapted version of the formula proposed by Paas and Van Merriënboer (1993). According to the formula: Efficiency (E) = $([P - T] / \sqrt{2})$, in which the P is standardized scores of post-test performance and T is standardized learning time. This formula is based on the assumption that participants who spend less learning time to achieve higher post-test performance have a high-learning efficiency, and participants who spend more learning time to achieve lower post-test performance have a low-learning efficiency.

Procedures

The experiment was conducted in a group setting in a computer classroom in which the participants individually learned with the corresponding narrated slideshow. The experiment had three phases: preparing and prior knowledge test, intervention, and post-test. In the preparing and prior knowledge test phase, participants were briefed about the study purpose, filled in their background information, and completed the prior knowledge test. In the intervention phase, participants were assigned to the three conditions and completed learning tasks respectively (please see the details of the learning materials for each

condition in the Learning materials section). The mental effort rating scale was applied immediately after learning. Then the participants completed post-tests in the third phase immediately after the intervention phase.

Results

An analysis of variance (ANOVA) to examine the differences in prior knowledge between the three conditions revealed no significant differences, $F(2, 93)=1.36, p>.05$. Table 2 lists the mean scores and standard deviations for the dependent variables including the post-test score, cognitive load and learning efficiency. Differences between conditions for each of these measures were analyzed with ANOVAs, and followed up with planned contrasts based on the hypotheses. The statistical significance for all tests was set at $p<.05$. The results of the ANOVAs revealed significant differences in the post-test scores, $F(2, 93)=3.42, MSE=11.61, p=.037, \eta_p^2=.07$, and learning efficiency, $F(2, 93)=61.37, MSE=.48, p<.001, \eta_p^2=.57$ but not in cognitive load, $F(2, 93)=1.21, MSE=2.59, p>.05$.

CPK condition vs. SP condition

The results of planned contrasts showed that the participants in the CPK condition performed better than the participants in the SP condition on the post-test ($t(93)=1.96, p=.03$, Cohen's $d=.51$) and learning efficiency ($t(93)=10.55, p<.001$, Cohen's $d=2.73$). No difference was found in cognitive load in learning phase ($t(93)=-.23, p>.05$).

CPW condition vs. SP condition

The results of planned contrasts showed that there were no significant differences between the CPW condition and SP condition for the post-test scores ($t(93)=-.51, p>.05$) and cognitive load in the learning phase ($t(93)=-1.45, p>.05$). Regarding learning efficiency, participants in the CPW condition performed better than participants in the SP condition ($t(93)=8.21, p<.001$, Cohen's $d=1.87$).

CPK condition vs. CPW condition

The results of planned contrasts revealed that participants in the CPK condition performed better than participants in the CPW condition in the post-test ($t(93)=2.49, p=.01$, Cohen's

Table 2 Means and standard deviations for the variables

Variable	CPK $N=31$		CPW $N=33$		SP $N=32$	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Post-test (0–31)	24.97	2.66	22.85	3.56	23.28	3.85
Cognitive load in learning phase (1–9)	4.00	1.51	3.52	1.77	4.09	1.53
Learning efficiency	.76	.54	.33	.72	–1.07	.78

$d = .67$) and learning efficiency ($t(93) = 2.49, p = .01$, Cohen's $d = .68$). However, there was no significant difference in cognitive load in the learning phase ($t(93) = 1.20, p > .05$).

Discussion

This study examined the temporal contiguity effect in narrated slideshow by comparing three popular presentation formats with different synchronization level of visual information and corresponding narration. The results partially supported the first hypothesis that the two concurrent presentation formats (CPK and CPW) would yield better learning outcomes than the sequenced presentation format (SP). Participants in the CPK condition performed significantly better than participants in the SP condition, for both learning and learning efficiency; and participants in the CPW condition performed better than participants in the SP condition only in terms of learning efficiency.

The analysis results showed that there was no difference in the post-test scores between the CPW and SP conditions. One possible reason for this finding is that participants in the SP condition had a longer learning time than participants in the other two concurrent presentation conditions. That is, even though learners in the SP condition needed to keep narration in working memory for a long time, the extra time for the presentation of corresponding visual information allowed them to search and match visual information which remained visible on the slide. In addition, as the visual information was presented at once, participants in the CPW condition still needed to hold narration in working memory for searching the elements on the slide which is referred to in the narration, even though the duration for holding was shorter than in the SP condition. Accordingly, the mentioned visual search processes might cost unnecessary cognitive resources and impede learning in both the CPW and SP conditions.

The temporal contiguity effect is a generally acknowledged effective instructional design principle for narrated animation (e.g., Mayer & Anderson, 1991; Mayer & Sims, 1994; Mayer et al., 1999). However, the results of the current study showed that this effect was only applicable in the narrated slideshow, in which all visual information was presented one by one, but was not applicable in all visual information was presented at once.

The results supported the second hypothesis that the concurrent presentation-key point format would lead to better learning outcomes than the concurrent presentation-whole format. Participants in the CPK condition showed better post-test scores and learning efficiency than participants in the CPW condition. A possible reason is that presenting key points may be a better way to reduce visual search by directing learners' attention to specific visual information compared with showing all information at once.

Directing attention, a key issue in multimedia learning (e.g., Crooks et al., 2012), is especially important in the design of narrated slideshow, where narration is dynamic and transient, but visual information is static (Kalyuga, 2012). For the CPW condition in which all visual information was concurrently presented as a whole, learners needed to keep narration in working memory and then search and match the corresponding visual information on the slide. This process may impose a heavy cognitive load on learners. Cueing or signaling is a method to draw learners' attention to essential elements of the learning materials (de Koning et al., 2009). Different types of cueing and signaling have been applied, such as arrow-line cueing for the integration of text and figure on a slide (Liu et al., 2013) and colors for the integration of figure with text or narration (Crooks et al., 2012). For example, in the experiment of Crooks et al. (2012), each specific part of the illustration temporarily changed color when

mentioned by text or narration and then changed back to its default color when the other part was mentioned.

For multimedia types like narrated slideshow, the results of this study pointed out that concurrent presentation-key point could be seen as a new format to direct learners' attention. Learners' attention will be directed to the related elements of the learning materials because the key points of visual information on a slide are presented at the moment the corresponding narration is mentioned, and remain present until the next slide is presented. In this case, the appearance of the key points provides a cue for what the learner should attend to. This feature makes learners' visual search and match with corresponding narration easier. Besides, different from traditional cueing/signaling that has been widely applied in designing animations (e.g., Crooks et al., 2012), key points of the visual information are still on a slide even when the corresponding narration ended and the next key point in the same slide is presented. To some degree, this feature could reduce the negative effect resulting from the transience of visual information.

The findings also support the explanation of the temporal contiguity effect based on dual-coding theory (Mayer & Anderson, 1991). For the CPK condition, the building of connections between the elements in the narration and the elements in visual information on the slide was facilitated because the elements of visual information were presented one by one on the slide together with corresponding narration.

Limitations of this study

Some limitations of this study should be considered. First, it should be noted that there were more narrative words in the sequenced version than in the concurrent versions, mainly because some words were used to connect the narration and the subsequent displayed visual information for sequenced version. In addition, as the participants of the current study were undergraduates, more research is needed to see whether the current results can be generalized to populations of a different age. Moreover, the gender distribution was not equal, because the number of female participants was larger than the number of male participants. The single-sample nature (homogeneity of gender) of the study might also be a reason for the nonsignificant post-test differences between the CPW and SP conditions (for a meta-analysis see, Castro-Alonso et al., 2019). Future research is recommended to replicate this study with more equal gender distributions. Finally, the analysis results showed that there was no difference in cognitive load among the three conditions. A possible reason for this is that only ten out of 27 slides of the narrated slideshow were designed for the treatment and perceptions of mental effort were measured after the participants had read all 27 slides. The perceived mental effort on the 10 treatment slides might be affected by the other 17 non-treatment slides. According to this, the differences between the perceptions of mental effort between the three conditions might not have been reflected in the ratings. Future research could try other measurement methods to dynamically track participants' mental efforts perceived from each slide or to increase the percentage of the treatment slides.

Suggestions for future studies

Future studies could replicate and further validate the results of this study through large-scale online experiments. In addition, this study cannot provide direct evidence in terms of cognitive processes to support claims regarding the reduction of learners' cognitive load by the different types of presentation designs. For example, more evidence is required to support the claim that the design of concurrent presentation-key points is favorable for learners to focus on the important parts of visual information, and the visual search and match with corresponding narration easier. Eye tracking technology could be used to uncover the cognitive processes involved in different designs of narrated slideshow. Eye tracking technology is a useful tool for measuring cognitive processes in research (Alemdag & Cagiltay, 2018; Liao et al., 2020; Lin et al., 2015) by enabling researchers to detect the allocation of attention (e.g., Hsu et al., 2019; Kokoc et al., 2020).

Conclusion

Narrated slideshow has been widely applied in digital learning environments. Different from narrated animation, in which both visual and verbal information are presented dynamically, narrated slideshow consists of a combination of static visual information and dynamic verbal information. The results of this study indicated that the temporal contiguity effect observed in the narrated animation can only be found in the narrated slideshow with visual information sequentially appeared by key points. The same effect cannot be observed in the narrated slideshow with all visual information presented at once. Based on the results, the current study provides a new point of view: when learning from a narrated slideshow (visual information is presented in a static way), the occurrence of a temporal contiguity effect depends not only on whether the presentations of narration and visual information are concurrent or not, but also on how concurrently they are presented. These findings extend the theoretical knowledge of the temporal contiguity effect and provide evidence-based guidelines for the design of narrated slideshow for instructional designers. Teachers are also encouraged to apply the new instructional guideline into their instructional practices.

Acknowledgements We would like to thank the editor of *Educational Technology Research and Development* and anonymous reviewers, who provided all the valuable comments and suggestions. Also, we would like to express our gratitude to the Ministry of Science and Technology (MOST) in Taiwan for its financial support to this research under Grant No. MOST 106-2511-S-003-033-MY3 and MOST 108-2511-H-003-044-MY4. Moreover, we sincerely appreciate the funding offered by the "Institute for Research Excellence in Learning Sciences" of National Taiwan Normal University (NTNU) from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan. Finally, we would like to thank all the people who helped us and supported this research. This research would not have been possible without them.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

- Alemdag, E., & Cagiltay, K. (2018). A systematic review of eye tracking research on multimedia learning. *Computers and Education*, 125, 413–428. <https://doi.org/10.1016/j.compedu.2018.06.023>
- Ayres, P., & Paas, F. (2007). Can the cognitive load approach make instructional animations more effective? *Applied Cognitive Psychology*, 21(6), 811–820. <https://doi.org/10.1002/acp.1351>
- Ayres, P., & Paas, F. (2009). Interdisciplinary perspectives inspiring a new generation of cognitive load research. *Educational Psychology Review*, 21, 1–9. <https://doi.org/10.1007/s10648-008-9090-7>
- Ayres, P., & Sweller, J. (2005). The split-attention principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 135–146). Cambridge University Press.
- Castro-Alonso, J. C., Wong, A., Adesope, O. O., Ayres, P., & Paas, F. (2019). Gender imbalance in instructional dynamic versus static visualizations: A meta-analysis. *Educational Psychology Review*, 31, 361–387. <https://doi.org/10.1007/s10648-019-09469-1>
- Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 3(3), 149–210. <https://doi.org/10.1007/BF01320076>
- Crooks, S. M., Cheon, J., Inan, F., Ari, F., & Flores, R. (2012). Modality and cueing in multimedia learning: Examining cognitive and perceptual explanations for the modality effect. *Computers in Human Behavior*, 28(3), 1063–1071. <https://doi.org/10.1016/j.chb.2012.01.010>
- De Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review*, 21, 113–140. <https://doi.org/10.1007/s10648-009-9098-7>
- DeLozier, S. J., & Rhodes, M. G. (2017). Flipped classrooms: A review of key ideas and recommendations for practice. *Educational Psychology Review*, 29(1), 141–151. <https://doi.org/10.1007/S10648-015-9356-9>
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. (2011). *How to design and evaluate research in education* (8th ed.). McGraw Hill.
- Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction*, 16(6), 511–525. <https://doi.org/10.1016/j.learninstruc.2006.10.001>
- Hsu, C. Y., Chiou, G. L., & Tsai, M. J. (2019). Visual behavior and self-efficacy of game playing: An eye movement analysis. *Interactive Learning Environments*, 27(7), 942–952. <https://doi.org/10.1080/10494820.2018.1504309>
- Kalyuga, S. (2009). *Managing cognitive load in adaptive multimedia learning*. Information Science Reference.
- Kalyuga, S. (2012). Instructional benefits of spoken words: A review of cognitive load factors. *Educational Research Review*, 7(2), 145–159. <https://doi.org/10.1016/j.edurev.2011.12.002>
- Khacharem, A., Trabelsi, K., Engel, F. A., Sperlich, B., & Kalyuga, S. (2020). The effects of temporal contiguity and expertise on acquisition of tactical movements. *Frontiers in Psychology*, 11, 413. <https://doi.org/10.3389/fpsyg.2020.00413>
- Kokoc, M., HIGaz, H., & Altun, A. (2020). Effects of sustained attention and video lecture types on learning performances. *Educational Technology Research and Development*, 68(6), 3015–3039. <https://doi.org/10.1007/s11423-020-09829-7>
- Liao, C. N., Chang, K. E., Huang, Y. C., & Sung, Y. T. (2020). Electronic storybook design, kindergartners' visual attention, and print awareness: An eye-tracking investigation. *Computers and Education*, 144, 103703. <https://doi.org/10.1016/j.compedu.2019.103703>
- Lin, Y. C., Liu, T. C., & Sweller, J. (2015). Improving the frame design of computer simulations for learning: Determining the primacy of the isolated elements or the transient information effects. *Computers and Education*, 88, 280–291. <https://doi.org/10.1016/j.compedu.2015.06.001>
- Liu, T. C., Lin, Y. C., & Paas, F. (2013). Effects of cues and real objects on learning in a mobile device supported environment. *British Journal of Educational Technology*, 44(3), 386–399. <https://doi.org/10.1111/j.1467-8535.2012.01331.x>
- Liu, T. C., Lin, Y. C., Hsu, C. Y., Hsu, C. Y., & Paas, F. (2021). Learning from animations and computer simulations: Modality and reverse modality effects. *British Journal of Educational Technology*, 52(1), 304–317. <https://doi.org/10.1111/bjet.12996>
- Liu, T. C., Lin, Y. C., & Kalyuga, S. (2022). Effects of complexity-determined system pausing on learning from multimedia presentations. *Australasian Journal of Educational Technology*, 38(1), 102–114. <https://doi.org/10.14742/ajet.7267>

- Mammarella, N., Fairfield, B., & Di Domenico, A. (2013). When spatial and temporal contiguities help the integration in working memory: "A multimedia learning" approach. *Learning and Individual Differences, 24*, 139–144.
- Mayer, R. E. (2017). Using multimedia for e-learning. *Journal of Computer Assisted Learning, 33*(5), 403–423. <https://doi.org/10.1111/jcal.12197>
- Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology, 83*(4), 484–490. <https://doi.org/10.1037/0022-0663.83.4.484>
- Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology, 84*(4), 444–452.
- Mayer, R. E., Moreno, R., Boire, M., & Vagge, S. (1999). Maximizing constructivist learning from multimedia communications by minimizing cognitive load. *Journal of Educational Psychology, 91*(4), 638–643. <https://doi.org/10.1037/0022-0663.91.4.638>
- Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning? *Journal of Educational Psychology, 86*(3), 389–401.
- Mulenga, E. M., & Marbán, J. M. (2020). Prospective teachers' online learning Mathematics activities in the age of COVID-19: A cluster analysis approach. *EURASIA Journal of Mathematics, Science and Technology Education, 16*(9), em1872. <https://doi.org/10.29333/ejmste/8345>
- Mutlu-Bayraktar, D., Cosgun, V., & Altan, T. (2019). Cognitive load in multimedia learning environments: A systematic review. *Computers and Education, 141*, 103618. <https://doi.org/10.1016/j.compedu.2019.103618>
- Owens, P., & Sweller, J. (2008). Cognitive load theory and music instruction. *Educational Psychology, 28*(1), 29–45. <https://doi.org/10.1080/01443410701369146>
- Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*(1), 63–71. https://doi.org/10.1207/S15326985EP3801_8
- Paas, F., & Van Merriënboer, J. J. G. (1993). The efficiency of instructional conditions: An approach to combine mental-effort and performance measures. *Human Factors, 35*(4), 737–743. <https://doi.org/10.1177/001872089303500412>
- Paas Jeroen, F. G. W. C., & Van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology, 86*(1), 122–133. <https://doi.org/10.1037/0022-0663.86.1.122>
- Schüler, A., Scheiter, K., Rummer, R., & Gerjets, P. (2012). Explaining the modality effect in multimedia learning: Is it due to a lack of temporal contiguity with written text and pictures? *Learning and Instruction, 22*(2), 92–102. <https://doi.org/10.1016/j.learninstruc.2011.08.001>
- Sweller, J. (2020). Cognitive load theory and educational technology. *Educational Technology Research and Development, 68*, 1–16. <https://doi.org/10.1007/s11423-019-09701-3>
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer.
- Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–296.
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*, 261–292. <https://doi.org/10.1007/s10648-019-09465-5>
- Zhuhadar, L., Kruk, S. R., & Daday, J. (2015). Semantically enriched massive open online courses (MOOCs) platform. *Computers in Human Behavior, 51*, 578–593. <https://doi.org/10.1016/j.chb.2015.02.067>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Tzu-Chien Liu is Professor of the Department of Educational Psychology and Counseling, and Institute for Research Excellence in Learning Sciences at National Taiwan Normal University. His research interests mainly focus on technology enhanced learning, learning and instruction, cognitive load theory, instructional design and digital literacy. He is now the associate editor of the Journal of Research in Education Sciences (Scopus index).

Yi-Chun Lin is Postdoctoral Research Fellow of the Department of Educational Psychology and Counseling at National Taiwan Normal University. She received her PhD in education from the National Central University in Taiwan. Her research interests include educational psychology, educational technology and learning and instruction.

Fred Paas is Professor of Educational Psychology at the Department of Psychology, Education & Child Studies at Erasmus University Rotterdam in the Netherlands, and Professorial Fellow at the School of Education/Early Start, University of Wollongong, Australia. Since 1990 he has been using the theoretical framework of cognitive load theory to investigate the instructional control of cognitive load in the training of complex cognitive tasks. He is the editor-in- chief of the journal *Educational Psychology Review*, and on the editorial board of several other renowned journals, such as the *Journal of Educational Psychology*. He is a fellow of the American Educational Research Association.