International Journal of Educational Technology and Learning
Vol. 19, No. 1, pp. 64-71, 2025
DOI: 10.55217/101.v18i3.1002

The influence of learner modality preferences on perceived mental effort and performance in self-paced instructional videos

Chun-Ying Chen

Center for General Education, National Taichung University of Education, Taiwan.
Email: cychen666@mail.ntcu.edu.tw

Abstract

Self-paced online learning environments frequently employ narrated and captioned videos, each grounded in distinct multimedia learning principles. Narrated videos reflect the modality principle, which favors spoken over written text, whereas captioned videos align with the transient information principle, which emphasizes the benefits of written input. These principles are in tension—one privileges auditory input while the other privileges visual-verbal input—and they map onto the processing tendencies of auditory versus read/write learners, creating potential inconsistencies in instructional design. This study investigated how learner modality preference moderates the effects of text modality on perceived mental effort and performance in video-based learning. A 2 × 2 between-subjects ANOVA (text modality: narrated vs. captioned; modality preference: auditory vs. read/write) revealed a significant interaction effect on perceived mental effort but no effects on recall or transfer performance. Follow-up analyses showed that text modality significantly influenced perceived mental effort among auditory learners, with narrated videos reducing effort relative to captioned videos, whereas no such effect was observed for read/write learners. The findings indicate that mismatched formats may increase perceived effort without impairing learning outcomes, likely because pacing-control mechanisms help learners manage cognitive demands. The differential effects observed between auditory and read/write learners provide insight into the application of multimedia principles and reinforce the view that instructional effectiveness is better explained by cognitive principles, such as modality and segmenting, rather than by tailoring materials to individual learner preferences.

Keywords:
Cognitive theory of multimedia learning
Instructional video
Learner modality preference
Modality principle
Transient information principle.

Publisher:
Scientific Publishing Institute

Received: *3 September 2025*
Revised: *16 September 2025*
Accepted: *19 September 2025*
Published: *25 September 2025*

Funding: This study received no specific financial support.
Institutional Review Board Statement: The study involved minimal risk and adhered to ethical guidelines for social science fieldwork. Formal approval from an Institutional Review Board was not required under the policies of National Taichung University of Education, Taiwan. Informed verbal consent was obtained from all participants, and all data were anonymized to ensure participant confidentiality.
Transparency: The author confirms that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
Competing Interests: The author declares that there are no conflicts of interests regarding the publication of this paper.

1. Introduction

The popularization of video-sharing platforms such as YouTube has established online videos as a primary learning resource, shifting attention toward instructional video design (Fyfield, Henderson, & Phillips, 2022; List, 2018). Instructional videos typically integrate dynamic visuals with verbal information delivered through spoken or written text (Fiorella & Mayer, 2018) reflecting a multimodal approach grounded in multimedia learning principles (Fyfield et al., 2022; Mayer & Fiorella, 2022).

Among these principles, the segmenting, modality, and transient information principles are frequently applied to instructional video design. While strong empirical support exists for the segmenting principle, evidence for others, such as the modality principle, remains mixed under varying conditions (Fyfield et al., 2022). Fyfield et al. (2022) further noted that most multimedia design principles were derived from studies using media other than videos, underscoring the need for research that validates these principles specifically in video-based learning contexts.

Instructional videos commonly adopt two text modalities: narrated (spoken explanations) and captioned (written text synchronized with visuals, typically displayed at the bottom of the screen). Each format aligns with different cognitive principles. Narrated videos embody the modality principle, which favors spoken over written words, whereas captioned videos reflect the transient information principle, which highlights the benefits of written text for reducing the fleeting nature of spoken input. However, these principles are in tension: one privileges auditory input, the other visual-verbal input. Similarly, the two formats may align differently with learner preferences, benefiting auditory versus read/write learners.

Given these inconsistencies, this study examined how learners with distinct modality preferences interact with narrated and captioned formats in self-paced instructional videos. The results contribute to a more nuanced understanding of the modality and transient information principles in the design of self-paced video-based learning environments.

2. Literature Review

2.1. Multimedia Design for Effective Instructional Videos

Multimedia design principles are grounded in the Cognitive Theory of Multimedia Learning (CTML), which emphasizes the working memory (WM) bottleneck as a key determinant of instructional effectiveness (Mayer, 2022). CTML posits that learning occurs through active cognitive processes—selecting, organizing, and integrating information—that operate across two distinct WM channels: a visual–pictorial channel and an auditory–verbal channel. Because WM has limited capacity and duration, learners can process only a restricted amount of information at any given time. Instructional effectiveness therefore depends on supporting these processes within the learner's available cognitive capacity, defined as the total processing resources across both channels.

Videos present information rapidly and through multiple modalities, making the management of cognitive processing essential to avoid overload. To address this, multimedia research has proposed several design principles, notably the segmenting, modality, and transient information principles. The segmenting principle suggests that learning improves when instruction is divided into learner-paced segments rather than presented as a continuous stream (Mayer & Fiorella, 2022). The modality principle advocates for spoken words over written text, since distributing information across both WM channels can reduce visual load and enhance learning compared to visual-only presentations (Castro-Alonso & Sweller, 2022). In contrast, the transient information principle favors written text, arguing that spoken words are fleeting and less likely to be fully processed compared to the more stable written form (Jiang & Sweller, 2022).

The segmenting principle has received strong empirical support, as learner-paced segmentation and pacing-control devices consistently reduce cognitive load and enhance performance (Fiorella, 2022; Fyfield et al., 2022). By contrast, the modality and transient information principles address different—and often conflicting—cognitive processing demands. While one promotes spoken over written words, the other privileges written over spoken. This theoretical tension highlights the need to examine how learners engage with different text modalities in video-based learning.

2.2. Learning in Different Text Modalities

Narrated videos reflect the modality principle, which emphasizes distributing information across auditory and visual channels to achieve more efficient processing than relying solely on the visual channel. The modality principle also addresses the split-attention effect (Sweller, Ayres, & Kalyuga, 2011) which arises when learners must integrate information from spatially separated sources. Under such conditions, learners may experience increased cognitive load because they can only attend to the visual demonstration or the written caption at a given moment, or must rapidly shift attention between them.

Captioned videos, in contrast, embody the transient information principle, which posits that permanent representations of information are superior to transient ones (Jiang & Sweller, 2022). Written captions reduce cognitive load because they can be re-read, unlike spoken narration, which is fleeting and may not remain in WM long enough for full comprehension (Jiang & Sweller, 2022; Leahy & Sweller, 2016). In addition, captions facilitate searching and skimming within instructional content, making pacing-control devices more effective with written text than with spoken narration (Chen & Yang, 2020; Chen & Yen, 2021). Captions may also encourage learners to replay video segments more frequently than narration, a behavior positively associated with learning outcomes (Chen & Yen, 2021; Schüler, Scheiter, & Gerjets, 2013).

Despite their advantages, these two formats embody contradictory design principles: the modality principle favors spoken over written words, whereas the transient information principle favors written over spoken. They also align differently with learner preferences, benefiting auditory versus read/write learners. Given these inconsistencies, examining how these learner types interact with different text modalities in video-based learning can deepen our understanding of the modality and transient information principles in instructional video design.

2.3. Learner Modality Preferences

Research on individual differences in learning has explored learning styles, learning preferences, and cognitive styles, all of which address how learners process information (Mayer & Massa, 2003). Despite widespread interest, there is no credible evidence that tailoring instruction to learning styles improves outcomes (Feldon, Jeong, & Clark, 2022). Nonetheless, the meshing hypothesis—the claim that learners perform better when instruction matches their preferred modality (e.g., auditory vs. visual text)—remains influential (Pashler, McDaniel, Rohrer, & Bjork, 2008) even though empirical support is lacking.

One complication is conceptual overlap and inconsistency. For instance, preferences may reinforce skill development in a modality through repeated practice (Lehmann & Seufert, 2020) yet stronger skills do not necessarily translate into superior learning outcomes (Clinton-Lisell & Litzinger, 2024). Similarly, the widely studied verbalizer–visualizer distinction is considered a cognitive style rather than a modality preference (Massa & Mayer, 2006) but its operationalization is problematic: "verbalizers" may rely on either written or spoken information, conflating read/write and auditory processing.

Empirical findings further challenge the meshing hypothesis. Studies with single-modality materials have consistently failed to demonstrate performance advantages for matched instruction (Rogowsky, Calhoun, & Tallal, 2015, 2020). Partial support comes from Lehmann and Seufert (2020) who found that visual learners benefited from text-based instruction, likely due to stronger text-processing strategies, while "auditive-ambiguous" learners performed equally well across modalities, suggesting flexibility rather than dependence on preference. Moreover, research on multimedia instruction indicates that instructional formats that optimize cognitive processing (e.g., segmentation, redundancy reduction) yield stronger effects than those tailored to learner preferences (Knoll, Otani, Skeel, & Van Horn, 2017; Kollöffel, 2012).

Taken together, these findings expose theoretical gaps and contradictions. While preferences may influence strategy use and subjective experience, they do not reliably predict learning performance. The blurred boundaries between preferences and cognitive styles, coupled with inconsistent definitions of constructs such as "verbalizers," further complicate the literature. These tensions highlight the need to examine how specific learner preferences—particularly auditory versus read/write—interact with instructional text modalities such as narration and captions in video-based learning.

2.4. Learners' Perceived Mental Effort During Learning

Paas and Van Merriënboer (1994) proposed a cognitive load model comprising three causal factors—task/environmental demands, learner characteristics, and their interaction—and three corresponding assessment factors: mental load, mental effort, and performance. Mental load reflects the demands imposed by the task or environment, independent of learner characteristics. Mental effort denotes the actual cognitive capacity allocated to a task, shaped by all three causal factors. Performance evaluates learning outcomes and likewise reflects their combined influence.

Paas and Van Merriënboer (1994) cognitive load model distinguishes between mental load (task demands independent of the learner), mental effort (the cognitive capacity actually invested in the task), and performance. Because this study examines how instructional modalities interact with learner preferences, mental effort was selected as the primary measure. Unlike mental load, it captures learner-dependent differences and reflects how learners allocate cognitive resources. Learners may compensate for higher task demands by investing additional effort to maintain performance (Sweller, Van Merrienboer, & Paas, 1998). Moreover, when instruction matches learner preferences, mental effort may be reduced through more efficient processing strategies (Lehmann & Seufert, 2020).

3. The Study

The question guided this research: Does learner modality preference (auditory vs. read/write) moderate the effect of instructional text modality (narrated vs. captioned videos) on learners' perceived mental effort and performance in self-paced video learning? Participants' subjective learning experiences were expected to vary depending on whether the instructional format matched their preferred modality. Specifically, learners may experience reduced mental effort with a matched format, due to effective processing strategies aligned with their preference (Lehmann & Seufert, 2020). Conversely, learners may perceive higher mental effort with an unmatched format, as ineffective modality-specific skills require additional cognitive resources to maintain performance. This dynamic reflects the interaction between task/environmental demands and learner characteristics, which can influence cognitive load (Paas & Van Merriënboer, 1994). Importantly, while modality preference may shape learners' subjective experience, it does not reliably predict objective performance outcomes (Knoll et al., 2017).

Accordingly, we hypothesized an interaction effect on perceived mental effort: auditory learners would report lower mental effort with narrated videos, and read/write learners would report lower mental effort with captioned videos. No interaction was expected for performance.

4. Method

4.1. Participants

A total of 78 university students (19–24 years old; 55 females, 23 males) were recruited from five introductory courses on digital literacy and technologies for non-computing majors at the Center of General Education in Taiwan. The experiment was conducted over two semesters in five regular classes, each with approximately 70–80 students. At the beginning of each class, students signed informed consent forms to participate in the data collection process. Following this, all participants completed a survey assessing their computer experiences and modality preferences using the VARK questionnaire.

The experiment was integrated into regular class activities. When all classes were over, we collected data from those who had never used Illustrator software and had a clear auditory or read/write preference. Each class yielded approximately 10-20 students meeting the sampling criteria, resulting in a final sample of 78 students (out of 360) for data analysis (see Table 1).

4.2. Materials

All materials were web-based, comprising the learning materials, a recall test and a transfer task to measure learning outcomes, a subjective cognitive load rating to assess perceived mental effort, and the VARK questionnaire to evaluate learner modality preferences. All materials were presented in traditional Chinese, the official language of Taiwan.

4.3. Learning Materials

The learning materials consisted of web-based instructional videos in narrated and captioned formats, adopted from Chen and Yang (2020). The videos taught participants how to use Adobe Illustrator's pen tool (Bezier tool), with content divided into five parts: (a) course outline, (b) pen tool basics, (c) layers and placing images, (d) beginning to trace, and (e) editing paths (Chen & Yang, 2020). This procedural computer task involved direct manipulation interface controls, which are visually dynamic and non-intuitive—for example, drawing lines or curves via drag-and-release mouse actions. The videos captured the computer screen while the instructor demonstrated each procedure.

The instructions applied the segmenting principle for effective learning (Chen & Yang, 2020). The content was divided into five labeled video segments, presented as a navigational menu on the left side of the screen, with the instructional content displayed on the right. Learner control was supported through: (a) the navigational menu, (b) stop/play buttons and a slide bar for fast-forwarding and rewinding, (c) a volume control for audio narration (disabled in captioned videos), and (d) a full-screen option. Both formats used the same navigational menu, learner-control features, and identical verbal explanations. The total duration of all video segments was 13:08 min (788 s), with an average of 3:17 min (197 s) per segment.

4.4. Recall Test and Transfer Task

Learning outcomes were assessed during the test phase using a recall test and a transfer task. The online recall test included five true-or-false and five multiple-choice questions, with one point awarded per correct answer, yielding a maximum score of 10. Questions were adapted from the Adobe certification examination (https://www.certiport.com/adobe) and translated from English into traditional Chinese.

The online transfer task required participants to draw a simple-contour image using Illustrator's pen tool without instructional guidance. The task involved 14 major components, with one point awarded for each correctly completed component, for a maximum score of 14. Participants were not permitted to access the instructions during the test phase.

4.5. Perceived Mental Effort During Learning

The subjective cognitive load rating scale (Paas, 1992) is a 9-point scale measuring participants' perceived mental effort during learning. After the learning phase, participants indicated how much effort they exerted to follow the instructions (1 = lowest and 9 = highest).

4.6. VARK Questionnaire

We focused on modality preferences, specifically auditory and read/write learners, due to their contrasting tendencies when processing narrated and captioned videos. The VARK (Visual, Aural, Read/Write, Kinesthetic) model (Fleming, 2001) identifies an individual's preferred instructional modality, making it suitable for this study. According to VARK, visual learners prefer visual representations (e.g., graphs, charts, diagrams); aural learners prefer auditory information (e.g., lectures, discussions); read/write learners prefer written text; and kinesthetic learners are multimodal, favoring experiential and practical learning.

The VARK questionnaire (https://vark-learn.com/the-vark-questionnaire/) contains 16 questions, each offering four response options corresponding to the four modalities. Responses are scored to classify participants as having very strong, strong, or mild single preferences, or a combination of multiple preferences. For example, a question might ask:

I want to learn how to take better photos. I would:

- Use examples of good and poor photos showing how to improve them.
- Use the written instructions about what to do.
- Use diagrams showing the camera and what each part does.
- Ask questions and talk about the camera and its features.

In this study, visual learners were excluded due to the high visual demands of the learning materials and task. We also excluded multimodal learners, targeting participants with clear, strong single preferences. To filter these participants, we avoided bipolar measures that only separate auditory and read/write preferences, instead relying on the VARK classification for self-reported learning preferences.

The VARK is widely used due to its face validity, simplicity, and ease of administration (Leite, Svinicki, & Shi, 2010). Reported reliability estimates for the subscales are adequate: 0.85 (visual), 0.82 (aural), 0.84 (read/write), and 0.77 (Kinesthetic). Leite et al. (2010) noted that allowing multiple responses can lead to many participants being classified as having multiple preferences. To address this, we selected participants with very strong or strong single auditory or read/write preferences and no or mild preference in other modalities. Only 21.7% (78 of 360) of examinees met these criteria.

4.7. Procedure

At the beginning of the semester, students were briefed on the experiment's objectives and procedures and informed of their right to withdraw at any time. They then completed a questionnaire evaluating prior computer experience and modality preferences using the VARK questionnaire.

The experiment was conducted during the sixth week of the semester, coinciding with the curriculum unit on using Illustrator's pen tool. Sessions took place in the computer laboratory where students regularly received computer-based instruction. Each participant was randomly assigned to one of two video conditions, working individually at a computer, with participants in the narrated video condition using headphones.

The experiment included a learning phase followed by a test phase. Time-on-task was controlled for both phases. The instructional video lasted 13:08 min, and participants had 30 min for the learning phase, allowing multiple viewings as needed. Immediately afterward, they completed the self-reported perceived mental effort scale, followed by 10 min for the recall test and 40 min for the transfer task.

5. Results

Data were analyzed using a 2 × 2 between-subjects analysis of variance (ANOVA), with text modality (narrated vs. captioned) and learner modality preference (auditory vs. read/write) as between-subjects factors. When significant interactions were observed, follow-up simple effects analyses were conducted. The dependent variables were recall performance, transfer performance, and perceived mental effort during learning. A significance threshold of $\alpha = 0.05$ was adopted for all statistical tests. Descriptive statistics for all variables are reported in Table 1.

Table 1. Group means for dependent variables (Standard deviations in parentheses).

	Narrated Video			Captioned Video		
	Auditory (n=22)	**Read/Write (n=21)**	**Total (n=43)**	**Auditory (n=19)**	**Read/Write (n=16)**	**Total (n=35)**
Recall test (0-10)	6.77 (1.80)	7.62 (1.20)	7.19 (1.58)	7.63 (1.57)	7.44 (1.79)	7.54 (1.65)
Transfer task (0-14)	11.23 (1.88)	12.10 (2.05)	11.65 (1.99)	11.42 (2.71)	12.06 (2.18)	11.71 (2.47)
Mental effort (1-9)	5.59 (1.50)	7.00 (1.05)	6.28 (1.47)	7.32 (1.34)	6.25 (2.05)	6.83 (1.76)

Table 2 reports the ANOVA results for all measures. For both recall and transfer performance, neither main effects nor interaction effects were observed. For perceived mental effort, no significant main effects emerged; however, the interaction between text modality and learner modality preference was significant, $F(1, 74) = 13.24$, $p = 0.001$, partial $\eta^2 = 0.152$. Follow-up simple effects analyses revealed that auditory-preference participants reported significantly lower mental effort when learning from narrated video (M = 5.59) compared to captioned video (M = 7.32), $F = 13.65$, $p < 0.001$, partial $\eta^2 = 0.156$. In contrast, read/write-preference participants reported slightly lower mental effort when learning from captioned video (M = 6.25) compared to narrated video (M = 7.00), but this difference was not statistically significant, $F = 2.30$, $p = 0.134$, partial $\eta^2 = 0.030$. These findings suggest that text modality significantly influenced perceived mental effort among auditory learners but not among read/write learners.

Table 2. Summary of ANOVA results for all the measures.

		df	F	p	partial η^2
Recall performance	Text modality	1, 74	0.862	0.356	0.012
	Modality preference	1, 74	0.799	0.374	0.011
	Interaction	1, 74	2.034	0.158	0.027
Transfer performance	Text modality	1, 74	0.026	0.874	0.000
	Modality preference	1, 74	2.242	0.139	0.029
	Interaction	1, 74	0.050	0.823	0.001
Mental effort	Text modality	1, 74	2.054	0.156	0.027
	Modality preference	1, 74	0.255	0.615	0.003
	Interaction	1, 74	13.240	0.001	0.152

6. Discussion

This study examined how learners' modality preferences (auditory vs. read/write) interact with instructional text modality (narrated vs. captioned videos) in self-paced learning environments. The findings demonstrated a significant interaction effect on perceived mental effort, but no corresponding effect on recall or transfer performance. Specifically, auditory learners reported reduced mental effort when learning from narrated videos compared to captioned videos, whereas read/write learners reported slightly lower mental effort with captioned videos, although the difference was not statistically significant.

There are several explanations for why the modality principle reduced perceived mental effort across both learner types. First, narrated videos facilitated dual-channel processing, thereby enhancing cognitive efficiency. This finding is consistent with the assertion that the modality principle is particularly effective when multimedia instruction involves dynamic visual materials (e.g., animations or videos) with high visual complexity and novice learners (Mayer & Fiorella, 2022). In this study, all participants were novices in the subject matter, and the instructional content demonstrated visually dynamic, non-intuitive interface procedures for drawing. Under such conditions, close attention to the demonstration was required, and narration likely helped both learner types mitigate split-attention effects. Second, both auditory and read/write learners could rely on pacing-control features (e.g., replaying) to manage the transience of spoken information. Third, the verbal explanations provided in the videos were relatively simple and procedural, unlike complex texts such as scientific passages that demand advanced reading strategies (Lehmann & Seufert, 2020). Consequently, processing narration did not impose excessive difficulty for either group.

This study suggests that the transient information principle may increase perceived mental effort among auditory learners, but not among read/write learners. Captioned videos can induce split attention between visual demonstrations and written text, posing greater challenges for auditory learners who typically benefit from dual-channel processing. Moreover, auditory learners may lack efficient strategies for processing text-based information, making captions more cognitively demanding. By contrast, read/write learners can process captions more fluently, switching between text and visuals with less difficulty, thereby reducing the split-attention effect. In addition, captions provide a permanent, re-readable resource that can be easily skimmed, offering further advantages for learners who favor text-based input.

The results revealed no main or interaction effects of modality preference and text modality on learning outcomes, providing further evidence against the meshing hypothesis, which posits that instruction matched to learner preferences enhances learning outcomes (Pashler et al., 2008). Instead, the study suggests that mismatched conditions may increase learners' perceived effort without impairing their performance, likely because pacing-control mechanisms (e.g., replaying or segment navigation) help learners manage cognitive demands. Consistent with prior research, modality preference may influence subjective learning experiences but does not reliably predict objective performance outcomes (Knoll et al., 2017; Rogowsky et al., 2020).

At the same time, narration and captions fulfill distinct functions in instructional videos. From a pedagogical perspective, narrated videos supplemented with optional captions may offer the greatest flexibility. Learners can rely on narration to support comprehension during initial learning and use captions for review, content search, or reference. Providing access to both modalities not only accommodates diverse learner needs but also fosters the development of multimodal processing skills. As online instruction continues to expand, such flexibility will be essential in helping learners adapt to varied instructional contexts and optimize their learning strategies.

7. Limitations

Several limitations should be acknowledged. First, the study was conducted in a classroom setting, where uncontrolled factors such as learners' motivation, prior knowledge, and general cognitive abilities may have influenced the results. Second, perceived mental effort was measured using a self-report scale, which, while widely used, may be subject to bias and lacks the precision of physiological measures. Third, the study relied on the VARK questionnaire to classify learners' modality preferences. Although widely adopted, the VARK has faced criticism regarding its validity and scoring procedures (Leite et al., 2010). To mitigate this concern, the study focused on participants with strong and clear single preferences, excluding visual and multimodal learners. Finally, the study focused on teaching graphic drawing tasks using Adobe Illustrator and examined a specific instructional video for learning procedural computer tasks. Future research should explore other learning tasks and environments to validate these findings.

8. Conclusion

This study contributes to a deeper understanding of how learner modality preference interacts with instructional video design. While modality preference influenced learners' perceived mental effort, it did not affect performance outcomes. These findings reinforce the view that instructional effectiveness is better explained by cognitive principles such as the modality and segmenting principles, rather than by matching instruction to learner preferences.

Future research should extend this line of inquiry to other subject domains, task types, and learner populations, while incorporating multiple measures of cognitive load. By doing so, researchers can better clarify the boundary conditions under which instructional video design optimally supports both learning performance and subjective experience.

References

Castro-Alonso, J. C., & Sweller, J. (2022). *The modality principle in multimedia learning. In R.E. Mayer, & L. Fiorella (Eds.), The Cambridge handbook of multimedia learning* (3rd ed.). Cambridge, UK: Cambridge University Press.

Chen, C.-Y., & Yang, Y.-H. (2020). Investigation of the effectiveness of common representational formats in online learner-paced software training materials. *Innovations in Education and Teaching International*, *57*(1), 97-108. https://doi.org/10.1080/14703297.2018.1511443

Chen, C.-Y., & Yen, P.-R. (2021). Learner control, segmenting, and modality effects in animated demonstrations used as the before-class instructions in the flipped classroom. *Interactive Learning Environments*, *29*(1), 44-58. https://doi.org/10.1080/10494820.2019.1572627

Clinton-Lisell, V., & Litzinger, C. (2024). Is it really a neuromyth? A meta-analysis of the learning styles matching hypothesis. *Frontiers in Psychology*, *15*, 1428732. https://dx.doi.org/10.3389/fpsyg.2024.1428732

Feldon, D. F., Jeong, S., & Clark, R. E. (2022). *Fifteen common but questionable principles of multimedia learning. In R.E. Mayer, & L. Fiorella (Eds.), The Cambridge handbook of multimedia learning* (3rd ed.). Cambridge, UK: Cambridge University Press.

Fiorella, L. (2022). *Multimedia learning with instructional video. In R.E. Mayer, & L. Fiorella (Eds.), The Cambridge handbook of multimedia learning* (3rd ed.). Cambridge, UK: Cambridge University Press.

Fiorella, L., & Mayer, R. E. (2018). What works and doesn't work with instructional video. *Computers in Human Behavior*, *89*, 465-470. https://dx.doi.org/10.1016/j.chb.2018.07.015

Fleming, N. D. (2001). *Teaching and learning styles: VARK strategies*. Christchurch, New Zealand: N.D. Fleming.

Fyfield, M., Henderson, M., & Phillips, M. (2022). Improving instructional video design: A systematic review. *Australasian Journal of Educational Technology*, *38*(3), 155-183. https://doi.org/10.14742/ajet.7296

Jiang, D., & Sweller, J. (2022). *The transient information principle in multimedia learning. In R.E. Mayer, & L. Fiorella (Eds.), The Cambridge handbook of multimedia learning* (3rd ed.). Cambridge, UK: Cambridge University Press.

Knoll, A. R., Otani, H., Skeel, R. L., & Van Horn, K. R. (2017). Learning style, judgements of learning, and learning of verbal and visual information. *British Journal of Psychology*, *108*(3), 544-563. https://doi.org/10.1111/bjop.12214

Kollöffel, B. (2012). Exploring the relation between visualizer–verbalizer cognitive styles and performance with visual or verbal learning material. *Computers & Education*, *58*(2), 697-706. https://doi.org/10.1016/j.compedu.2011.09.016

Leahy, W., & Sweller, J. (2016). Cognitive load theory and the effects of transient information on the modality effect. *Instructional science*, *44*, 107-123. https://doi.org/10.1007/s11251-015-9362-9

Lehmann, J., & Seufert, T. (2020). The interaction between text modality and the learner's modality preference influences comprehension and cognitive load. *Frontiers in Psychology*, *10*, 2820. https://doi.org/10.3389/fpsyg.2019.02820

Leite, W. L., Svinicki, M., & Shi, Y. (2010). Attempted validation of the scores of the VARK: Learning styles inventory with multitrait–multimethod confirmatory factor analysis models. *Educational and Psychological Measurement*, *70*(2), 323-339. https://doi.org/10.1177/0013164409344507

List, A. (2018). Exploring the impact of video-sharing platforms on modern learning. *Journal of Educational Technology*, *12*(3), 45–60.

Massa, L. J., & Mayer, R. E. (2006). Testing the ATI hypothesis: Should multimedia instruction accommodate verbalizer-visualizer cognitive style? *Learning and Individual Differences*, *16*(4), 321-335. https://doi.org/10.1016/j.lindif.2006.10.001

Mayer, R. E. (2022). *Cognitive theory of multimedia learning. In R.E. Mayer, & L. Fiorella (Eds.), The Cambridge handbook of multimedia learning* (3rd ed.). Cambridge, UK: Cambridge University Press.

Mayer, R. E., & Fiorella, L. (2022). *Principles for managing essential processing in multimedia learning. In R.E. Mayer, & L. Fiorella (Eds.), The Cambridge handbook of multimedia learning* (3rd ed.). Cambridge, UK: Cambridge University Press.

Mayer, R. E., & Massa, L. J. (2003). Three facets of visual and verbal learners: Cognitive ability, cognitive style, and learning preference. *Journal of Educational Psychology, 95*(4), 833. https://doi.org/10.1037/0022-0663.95.4.833

Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*(4), 429-434. https://doi.org/10.1037/0022-0663.84.4.429

Paas, F. G., & Van Merriënboer, J. J. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review, 6*, 351-371. https://doi.org/10.1007/BF02213420

Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest, 9*(3), 105-119. https://doi.org/10.1111/j.1539-6053.2009.01038.x

Rogowsky, B. A., Calhoun, B. M., & Tallal, P. (2015). Matching learning style to instructional method: Effects on comprehension. *Journal of Educational Psychology, 107*(1), 64-78. https://doi.org/10.1037/a0037478

Rogowsky, B. A., Calhoun, B. M., & Tallal, P. (2020). Providing instruction based on students' learning style preferences does not improve learning. *Frontiers in Psychology, 11*, 511773. https://doi.org/10.3389/fpsyg.2020.00164

Schüler, A., Scheiter, K., & Gerjets, P. (2013). Is spoken text always better? Investigating the modality and redundancy effect with longer text presentation. *Computers in Human Behavior, 29*(4), 1590-1601. https://doi.org/10.1016/j.chb.2013.01.047

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.

Sweller, J., Van Merrienboer, J. J. G., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296. https://doi.org/10.1023/A:1022193728205