



Different types of textual cues in educational animations: Effect on science learning outcomes, cognitive load, and self-efficacy among elementary students

Lei Du¹ · Xiaoyu Tang¹ · Jingying Wang¹

Received: 3 December 2023 / Accepted: 25 July 2024

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

Abstract

Educational animation, recognized for its potential accessibility and engaging qualities, has become increasingly prevalent in classroom instruction. However, not all educational animations exhibit high quality or significantly enhance learning outcomes. This study addresses the need for optimizing educational animation design to enhance student learning outcomes and experiences, employing the construction-integration model. We developed three types of educational animations: subtitled textual cue (STC), keyword textual cue (KTC), and structured textual cue (CTC). Through a quasi-experimental research design, 257 fifth-grade students were assigned to three groups, each exposed to one type of textual cue. The results indicate that CTC leads to superior achievement, knowledge retention, higher self-efficacy, and the lowest cognitive load. In comparison, KTC demonstrates moderate results, while STC yields the poorest outcomes. Furthermore, there is a significant negative correlation between achievement and cognitive load, and a significant positive correlation between achievement and self-efficacy. Additionally, there is a significant positive correlation between the "faded effect" of knowledge retention and self-efficacy. These findings highlight the superior learning outcomes and experiences associated with CTC. Based on these findings, recommendations are provided for future educational animation design and instructional practices.

Keywords Educational animation · Self-efficacy · Cognitive load · Learning outcomes · Multimedia learning

✉ Jingying Wang
wangjingying8018@126.com

Lei Du
dulei0196@163.com

Xiaoyu Tang
xiaoyut@mail.bnu.edu.cn

¹ Faculty of Education, Beijing Normal University, Beijing, P. R. China

1 Introduction

Educational animation, a pedagogical tool amalgamating audio-visual elements, exhibits significant appeal and instructional potential in contemporary education. It is extensively employed across diverse learning environments to foster active student engagement and facilitate profound comprehension (Krieglstein et al., 2023), particularly emerging as a pivotal instrument in science education and related domains (Barak et al., 2011; Castro-Alonso et al., 2019). Concurrently, to enhance students' learning outcomes, textual cues are commonly integrated into educational animations to direct attention (Arslan-Ari et al., 2020; Lowe & Schnotz, 2014). However, despite the evident benefits of employing textual cues in educational animations to enhance learning experiences and comprehension, certain studies have highlighted that inadequately designed textual cues may diminish instructional successfulness by increasing cognitive load (Alpizar et al., 2020; Castro-Alonso et al., 2018; De Koning & Jarodzka, 2017).

To optimize future design and teaching practices in educational animation, thereby enhancing student learning outcomes and experiences, this study applied Kintsch's (1998) construction-integration model theory to develop three types of textual cues: subtitle textual cue (STC), keyword textual cue (KTC), and structured textual cue (CTC). The study specifically focuses on digital natives born after 2010, whose interaction and cognition with multimedia resources may differ significantly from previous generations (Tarchi et al., 2021). Employing a quasi-experimental approach, the study investigated how these different textual cues impact students' learning outcomes, including achievement and knowledge retention. Additionally, the study examined the contributions of educational animations with these textual cues on students' cognitive load and self-efficacy, while exploring the relationships among achievement, knowledge retention, self-efficacy, and cognitive load.

2 Literature review

2.1 Educational animation

Educational animation is a form of animated media designed specifically to teach or convey educational content. It utilizes visual storytelling, motion graphics, and often sound to simplify and illustrate complex concepts, making them easier to understand and more engaging for learners. (Krieglstein et al., 2023). Educational animations have five key characteristics, as summarized by Ploetzner and Lowe (2012) through a systematic literature review: (1) The technological platform for presenting animation is the electronic devices; (2) Educational animation aims to provide explicit explanations of the entities, structures, and processes involved in the subject to be learned; (3) Educational animation consists of entities created manually through drawing or other modeling techniques; (4)

Educational animation mimics the process of a series of different images to be learned. With advancements in technology, the use of educational animation in technology-enhanced learning environments has become more prevalent and easier to implement. (Castro-Alonso et al., 2019; Yilmaz, 2023). It helps students understand abstract concepts through dynamic visual presentations (Barak et al., 2011). Numerous studies highlight the significant potential of educational animation compared to other instructional media. For example, Ploetzner et al. (2021) found that educational animation is more beneficial than static visualization for supporting dynamic perceptual learning and the acquisition of kinematic models. In the field of science education, comprehensive educational practices have been developed by many scholars (Arslan-Ari et al., 2020). Türkay (2016) demonstrated that educational animations in physics classrooms positively affect students' knowledge retention and classroom engagement. Karlsson (2010) found that educational animations can engage students in collaborative problem-solving, leading to meaningful construction of scientific knowledge. Consequently, the use of educational animation in science learning is increasingly recommended by researchers (Tosun, 2022).

2.2 Cognitive load in educational animation

While the use of educational animation in pedagogy has gained widespread acceptance, some studies express concerns regarding its potential to impose higher cognitive loads (Castro-Alonso et al., 2018; Krieglstein et al., 2023). Cognitive load refers to the mental workload imposed on an individual's cognitive system during task processing or information acquisition. According to Paas and Van Merriënboer (1994), cognitive load can be divided into two components: mental load and mental effort. Mental load pertains to the demands placed on the cognitive system by the task itself. It is influenced by the complexity of the content and the manner in which information is presented. Mental effort refers to the amount of cognitive capacity that an individual allocates to meet the demands of the task. According to the Cognitive Theory of Multimedia Learning (Mayer, 2024), students must rapidly process information from both visual and auditory channels within a short period, when watching educational animations. They also need to integrate new information with prior knowledge and keep relevant information active in working memory (Mayer & Pilegard, 2014), all of which contribute to increased cognitive load. Research on cognitive load in educational animations highlights the need for well-designed animations. For example, poorly designed educational animations can overwhelm students' limited working memory capacity, leading to cognitive overload and hindering learning outcomes (De Koning & Jarodzka, 2017). In response, instructional designers must consider the constraints of cognition to optimize the use of educational animations. Consequently, there is a growing call in the scholarly community to enhance technological elements in educational animations and combine novel theoretical frameworks to reduce cognitive load and amplify the educational benefits (Arslan-Ari et al., 2020; Lowe & Schnotz, 2014).

2.3 Using textual cues in educational animations

The VARK theory highlights the diversity in learning styles, emphasizing the importance of offering multiple avenues for information delivery to meet the needs of all learners (Nguyen et al., 2024). This theory supports the shift from relying solely on visual and auditory stimuli in educational animations to integrating textual cues, representing a significant advancement in instructional design. By presenting textual information alongside animation, educators can create a multimodal learning environment that caters to diverse learning preferences and optimizes cognitive processing (Yilmaz, 2023). Textual cues act as guiding markers within the animation, directing students' attention to key instructional points and facilitating learning achievement (Clark & Mayer, 2016; Mayer & Pilegard, 2014; Wang et al., 2020). Textual cues help students corresponding information from the different external representations into a coherent mental representation to reduce students' cognitive load (Richter et al., 2018). Additionally, textual cues can improve the accuracy of knowledge retention (Rop et al., 2018). However, not all textual cues lead to better learning outcomes. Some researchers have demonstrated that inappropriate textual cues can result in a heavy cognitive load for students, negatively affecting their achievement and knowledge retention (Alpizar et al., 2020; Arslan-Ari et al., 2020). Berney and Betrancourt (2016) concluded through a meta-analysis that the pedagogical value of educational animations is affected by the manner in which accompanying textual information is presented. This suggests that textual cues in educational animations need to be carefully designed to be impactful and engaging, ensuring they enhance comprehension and retention.

2.4 Construction-integration model

In the realm of text information type and design, the Construction-Integration Model (CIM) posited by Kintsch (1998) serves as an important theoretical framework within cognitive science, elucidating the mechanisms underlying information processing during textual comprehension. The model emphasizes that readers' understanding and retention of text involve hierarchical representation and processing stages to construct coherent meaning, highlighting the dynamic and interactive nature of this construction process. CIM categorizes text understanding aimed at cognitive construction into three levels of representation: surface level, textbase level, and situation model level.

The surface level involves the decoding of textual information in the flow of discourse, akin to the supportive role played by subtitles in educational animations (Tarchi et al., 2021). Subtitled text information in educational animations refers to written language that appears synchronously with the animated content, providing viewers with concurrent written language cues while engaging with dynamic visual imagery (e.g., Clark & Mayer, 2016; Matthew, 2020). The textbase level corresponds to the conceptual and propositional content of the text's meaning, equivalent to critical information presented alongside the narrative in educational animations. Textual

cues can be subtle and flexible, utilizing color in written texts or varying intonation in spoken texts to draw attention to key terms (e.g., Rop et al., 2018; Wang et al., 2020). The situation model level is constructed through various forms of reasoning, including knowledge-based inference. During the integration process, students integrate relevant background knowledge activated by textual cues to generate a coherent representation (Kintsch, 2019). Moreover, it has been suggested that dynamic models can be constructed at finer-grained levels, illustrating how the knowledge structures depicted in animations scaffold students' partial cognitive structures (De Koning et al., 2009). However, despite these considerations, the application of situation model level informed textual cues within the domain of educational animations remains underexplored, prompting calls for empirical research from scholars such as Tarchi et al. (2021). Thus, this study adopts the principles of the CIM to develop educational animations featuring three distinct text types and explores the practical instructional effects of educational animations produced using different types of textual cues.

2.5 Self-efficacy and learning outcomes

Self-efficacy refers to an individual's belief in their ability to perform a specific task proficiently and confidently (Pintrich et al., 1991), which is crucial for motivation and perseverance in educational settings. In the context of multimedia instruction, particularly in learning environments enhanced by educational animations, self-efficacy is considered a crucial predictor of learning outcomes (Mayer & Pilegard, 2014; Semilarski et al., 2022). This assertion finds robust corroboration in a vast corpus of research, which underscores the advantageous influence of self-efficacy on academic accomplishments, particularly within the realm of science education (Lin, 2021; Tosun, 2022). The notion of self-efficacy as a favorable contributor to learning is further fortified by a recent systematic review of literature on technology-assisted education (Granic, 2022). Nonetheless, amidst these affirmations, social psychological research raises concerns. Researchers have drawn attention to potentially adverse consequences resulting from exaggerated self-efficacy, epitomized by the Dunning-Kruger effect, where individuals might overestimate their competence to meet set performance targets (Dunning, 2011). Given the pivotal role of self-efficacy in shaping scientific learning outcomes and its multifaceted impact on instructional quality, there is a need to explore the association between students' self-efficacy and achievement when exposed to diverse types of textual cues embedded in educational animations. This study endeavors to shed light on how varying textual cues might influence the self-efficacy-learning outcome relationship among the learners, thus providing insights into optimizing the integration of educational animations in science education to improve their learning benefits.

2.6 Present study

Educational animations are prevalent in teaching due to their vividness, intuitiveness, and simplicity, demonstrating great potential in improving learning outcomes

(Kriegelstein et al., 2023). However, the efficacy of educational animations depends not only on their visual and auditory appeal but also on their impact on learners' cognitive processing and cognitive load (Castro-Alonso et al., 2018; Paas & Van Merriënboer, 1994). Specifically, the strategic use of textual cues within animations can significantly aid learners in comprehending and retaining key concepts while reducing unnecessary cognitive load (Alpizar et al., 2020; Arslan-Ari et al., 2020; Richter et al., 2018). The question of which types of textual cues can best achieve these goals still requires further exploration. Moreover, self-efficacy has been identified as a crucial factor influencing learning effectiveness (Pintrich et al., 1991; Mayer & Pilegard, 2014; Seminarski et al., 2022). Interestingly, high self-efficacy in multimedia environments has been found to potentially harm learning outcomes (Dunning, 2011). Whether self-efficacy in educational animations with different types of textual cues also exhibits this double-edged effect is worth investigation.

Despite the importance of these factors, research on the effects of different types of textual cues in educational animations on learning effectiveness, cognitive load, and self-efficacy remains limited. Therefore, this study, grounded in the Construction-Integration model (Kintsch, 1998), developed three types of textual cues for educational animations: subtitled textual cue (STC), keyword textual cue (KTC), and structured textual cue (CTC). A quasi-experimental design was employed to compare the teaching effects of these three types of textual cues on learning performance and psychological experience. This study aims to fill gaps in the existing literature, provide a scientific basis for the design of educational animations, and offer practical guidance to educators on balancing self-efficacy and actual learning needs in educational design.

Specific research questions include:

RQ1: What differences exist in the achievement of students when exposed to educational animations with different types of textual cues?

RQ2: What differences exist in the knowledge retention of students when exposed to educational animations with different types of textual cues?

RQ3: What differences exist in the cognitive load experienced by students when exposed to educational animations with different types of textual cues?

RQ4: What differences exist in the self-efficacy of students when exposed to educational animations with different types of textual cues?

RQ5: What correlations exist between achievement, knowledge retention, cognitive load, and self-efficacy of students when exposed to educational animations?

3 Method

3.1 Participant

The study recruited a sample consisting of fifth-grade pupils from a Chinese public primary school, with ages ranging from 10 to 11 years old. A total of 261 students volunteered to take part in the experiment, with four unable to complete the entire study because of sickness or scheduling issues, resulting in 257 valid

subjects overall. All participating students and their parents were informed that the data collected would be anonymized and solely used for research purposes. All participants in the study shared the same science teacher, who provided full cooperation throughout the experimental process, thereby guaranteeing a standardized teaching tempo and consistent content delivery across the board.

3.2 Research procedure

The research procedure is shown as Fig. 1. During the preparation phase, we developed three types of animated materials required for the study and recruited participants. All participants were informed about the experiment’s basic information and procedure without compromising the research’s integrity. Subsequently, the researchers randomly assigned all participants into three groups. The STC group consisted of 87 participants, with 52% male and 48% female. The KTC group had 84 participants, with 49% male and 51% female. The CTC group comprised 86 participants, with 51% male and 49% female.

Based on the study by Arslan-Ari et al. (2020) on the effects of prior knowledge on learning outcomes from educational animations, all participants completed a prior knowledge test. A one-way ANOVA indicated similar levels of prior knowledge among the three groups ($F = 1.021, p > 0.05$). Additionally, the subject of this study was "Earth’s rotation", and none of the participants had

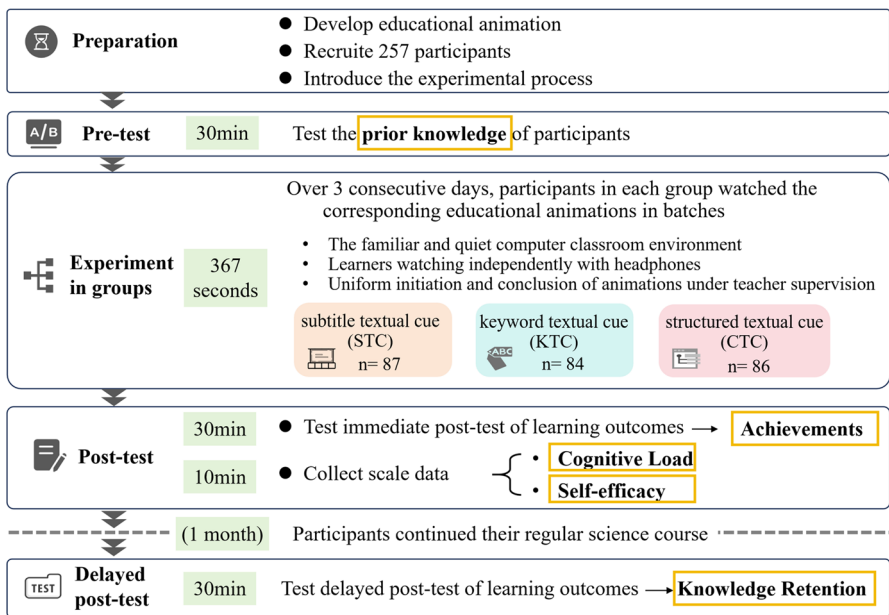


Fig. 1 Research procedure

previously studied the relevant course. Therefore, considering these two points, it can be concluded that prior knowledge is unlikely to interfere with the results of this experiment.

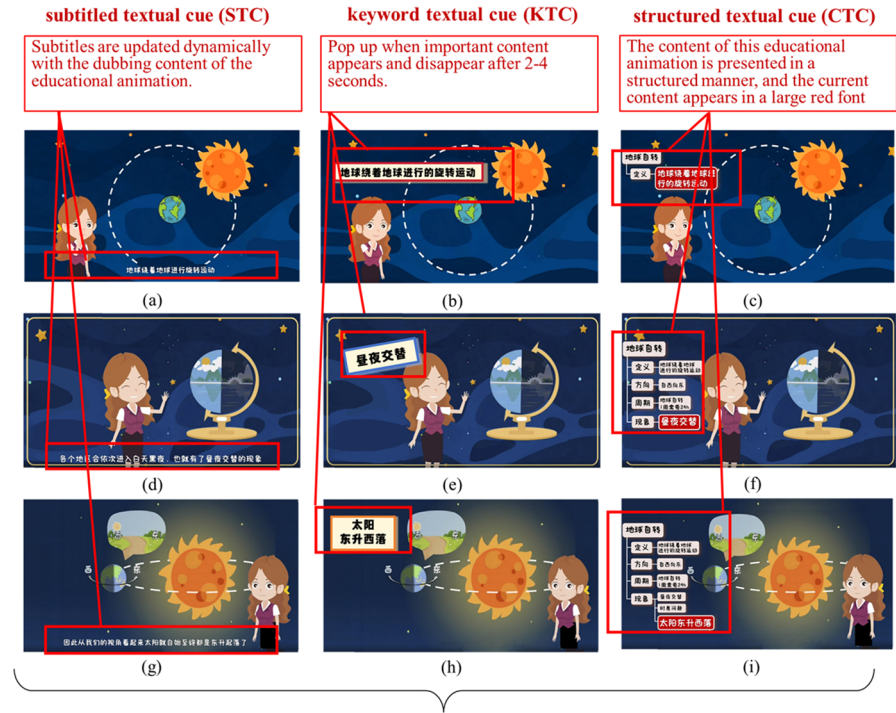
Over three consecutive days, participants in each group watched the corresponding educational animations in batches, with each type of animation lasting 367 s. To minimize potential distractions and other sources of interference, participants viewed the educational animations in a familiar computer classroom environment. In this learning environment, the computer systems operated under the centralized supervision of a teacher. The teacher initiated and concluded the animations uniformly, standardizing the collective viewing process. Throughout the animation, the playback speed and video progress were predefined and locked, preventing individual learners from adjusting the speed or progress in real-time. Each student was provided with headphones to watch the videos independently. During the viewing, students remained in a quiet environment, which helped them focus and engage with the presented material.

After watching the educational animations, students had a five-minute break, followed by a 30-min knowledge post-test and a combined 10-min cognitive load and self-efficacy questionnaire to assess the achievement and learning experiences of the three groups.

One month later, all participants from the three groups took a 30-min delayed post-test at a unified time to evaluate their knowledge retention. It is important to note that during this month, all participants continued their regular science classes under the same science teacher, which did not involve the use of the three types of educational animations or content related to Earth's rotation.

3.3 Educational animation materials

The educational animation utilized in this study focused on the topic of "Earth's rotation". According to the national science curriculum standards of China, this topic is included in the curriculum for students in grades 5 or 6 (MOE, 2022). To ensure the validity of the content, the animation was adapted from current Chinese textbooks and expanded to meet the needs of the students. The development of educational animation materials adhered to established principles: (1) Utilize animation technology to simulate the Earth's rotation and the resulting shift between day and night, effectively presenting dynamic information that is not easily captured in static material; (2) Employ a visual design that aligns with the educational objective to engage students' curiosity, including vibrant hues, appealing interactive characters, and lucid scenes; (3) Divide intricate concepts into manageable, sequential steps and vividly transmit knowledge; (4) The teaching objectives and learning orientations are clearly defined, and the content is concise and presented in easily understandable language; (5) Interactive elements such as questioning are incorporated to increase student participation and interest in learning, as well as to promote metacognition and knowledge construction; (6) Both the content and presentation are appropriate for learners' cognitive levels; (7) The animated dialogue is accurately and dubbed with a moderate speech speed for full student comprehension.



Screenshots of the educational animation of the three types of textual cues at the same moment. This part explains how the sun rises in the east and sets in the west as the Earth rotates.

Fig. 2 The example screenshot of each textual cue

Three versions of textual cues were developed to correspond to the three design sets in this study, as presented in Fig. 2. The first type is the Subtitle textual cue (STC), as illustrated in (a) (d) (g). The STC adds subtitles at the bottom of the educational animation, which comprises all the dialogue text and is dynamically updated. To ensure visibility for students, the subtitle color is in contrast with the page (for example, white subtitles are used on a dark blue background). The second type is the Keyword Text Clues (KTC), as depicted in (b) (e) (h). KTC displays crucial keywords at relevant positions during the course content, and each keyword remains visible for 2–4 s before disappearing. To ensure keyword clarity, a large and prominent font is used. The third type is the structured textual cue (CTC), as depicted in (c) (f) (i). Structured textual cues are dynamically generated following the development of the animated content, always displaying on the left screen. Newly appearing textual cues are in red and slightly larger than the rest of the text. It is noteworthy that the educational animations viewed by the three groups were identical in content, with the only variation being the type of textual cues used. Furthermore, the duration of all educational animations was standardized to 367 s. The specific definitions, characteristics and uses of textual cues are shown in Table 1.

Table 1 The specific definitions, characteristics and uses of subtitle textual cue (STC), keyword textual cue (KTC), structured textual cue (CTC)

Textual Cue Type	Levels of text understanding based on CIM	Definition	Function & Role	Features & Implementation Methods	Example
Subtitle textual cue (STC)	Surface level	Synchronous written language information provided at the bottom of animations that supplements and clarifies spoken dialogues to aid viewers' comprehension	<ol style="list-style-type: none"> 1. Reinforce memory and understanding of verbal content (Matthew, 2020) 2. Ensure critical conversations are not overlooked (Lowe & Schnotz, 2014) 	<ol style="list-style-type: none"> 1. Updates in real-time to align with audio dialogue 2. Utilizes high contrast color schemes (e.g., white text on black background or vice versa) 3. Considers layout placement to minimize visual distractions (e.g., the position at the inferior aspect of the display screen) 	Figure 2 (a)(d)(g)
Keyword textual cue (KTC)	Textbase level	Highlighting key vocabulary at pivotal moments within educational animations, guiding students towards core concepts and propositions	<ol style="list-style-type: none"> 1. Emphasize salient knowledge elements to facilitate memorization and learning of crucial information (Wang et al., 2020) 2. Enhance student focus and attention (Rop et al., 2018) 	<ol style="list-style-type: none"> 1. Displayed in screen areas closely tied to animation content 2. Keywords remain visible for a brief but noticeable duration (e.g., 2–4 s) 3. Employing large, bold fonts and potential highlighting to emphasize keywords 	Figure 2 (b)(e)(h)

Table 1 (continued)

Textual Cue Type	Levels of text understanding based on CIM	Definition	Function & Role	Features & Implementation Methods	Example
Structured textual cue (CTC)	Situation model level	During the progression of animated sequences, dynamically generated structured textual information is employed to facilitate learners in constructing situational models and coherent cognitive frameworks	<ol style="list-style-type: none"> 1. Activate learners' prior knowledge (Kintsch, 2019) 2. Help form internal knowledge networks and facilitate understanding and application of complex concepts (Farchi et al., 2021) 	<ol style="list-style-type: none"> 1. Fixed presence on the left side of the screen for easy reference 2. Newly introduced content differentiated using conspicuous colors like red and larger font sizes 3. Unfolds progressively in line with the logical sequence of the animation content, tightly integrated with instructional objectives and knowledge structures 	Figure 2 (c)(f)(i)

3.4 Data collection tools

3.4.1 Learning outcomes

The learning outcome is accomplished through three academic-level tests. Firstly, all the questions related to the "Earth's rotation" were screened from the educational exams organized by the local government educational authorities from 2015 to 2022, which ensured the trustworthiness of the quality of the questions. The Educational authorities classifies each question into three grades: easy, medium and difficult. The questions were categorized and grouped according to the difficulty level of the questions published by the educational authorities, and three sets of questions were obtained with the same difficulty level. Each set contained 4 multiple-choice questions, 2 fill-in-the-blank questions, and 1 short answer question, totaling 10 points. Three experienced elementary school science teachers were invited to re-evaluate the three sets of questions, which were found to be of a good level of difficulty, and the three sets of questions were used in the pre-test, the immediate post-test, and the delayed post-test to determine the students' basic level, achievement, and knowledge retention, respectively. Two researchers, specializing in science education, independently assessed learning outcomes, unaware of students' identities or group affiliations. Inter-rater reliability exceeded 0.95, indicating high consistency in scoring. Rasch analysis was employed to measure the validity of the test items, obtaining $\text{infit}_{\text{MNSQ}}$ and $\text{outfit}_{\text{MNSQ}}$ statistics. The results indicated that all item values fell within the range of [0.7, 1.15], suggesting that the item difficulty levels were well-aligned with the participants' ability characteristics.

3.4.2 Cognitive load and self-efficacy

The employed scales consist of two sections: cognitive load (CL) and self-efficacy (SE). Responses to all scale items were rated using a 5-point Likert scale. The cognitive load (CL) measure is derived from the scale developed by Paas and Van Merriënboer (1994), featuring two dimensions: mental load (ML) and mental effort (ME). Specifically, mental load encompasses 5 items, and mental effort includes 3 items. This scale, following adaptation by Hwang et al. (2013), has been effectively implemented in the context of elementary natural science curriculum evaluations, demonstrating positive outcomes in its application. The SE scale was adapted from an instrument originally created by Pintrich et al. (1991), consisting of 8 items, and stands as a widely utilized psychological instrument aimed at evaluating students' self-efficacy beliefs within learning environments. The Cronbach's α coefficients for the CL, ML, ME, SE scales were 0.841, 0.835, 0.858, and 0.892. To ensure the validity of the instruments, we conducted both content validity and construct validity checks. Content validity was ensured through expert review, where three experienced educators evaluated the relevance and clarity of each item in the context of the study objectives. All questionnaire items were translated into Chinese by expert translators and then reviewed by science education researchers. A pilot test was conducted with 10 fifth-grade students from the same school who did not participate in the main experiment. Based on their feedback, adjustments were made to guiding

language and certain phrasings to ensure each question could be accurately comprehended by the students. Construct validity was examined using factor analysis. The Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy and Bartlett’s test of sphericity were employed to confirm the suitability of the data for factor analysis. The KMO value of CL, ML, ME, SE scales were 0.703, 0.717, 0.833 and 0.762. The Bartlett’s test of sphericity was significant ($p < 0.001$), indicating that the factor analysis was appropriate. The factor loadings for all items were above 0.5, confirming that the scales effectively measured their respective constructs.

3.5 Data analysis

The study examined the influence of three types of textual cues (STC, KTC, and CTC) on students’ learning outcomes following the viewing of educational animations, employing various statistical analytical methods. All data demonstrate normal distribution characteristics, and have satisfied the requirements of the following variance analysis’s assumption, as confirmed by the hypothesis test. Initially, repeated measures analysis of variance was used to analyze pretest, post-test, and delayed post-test data across the three groups to assess both immediate and delayed effects comprehensively. To address potential confounding effects of students’ prior knowledge levels, ANCOVA was employed, treating pretest scores as a covariate. The study compared differences in immediate learning effectiveness between groups. For investigating the impacts of knowledge retention and the "faded effect" on knowledge retention, one-way ANOVA was utilized to determine which types of textual cues were more beneficial for maintaining knowledge over the long term and to elucidate distinctions in their contributions. To examine cognitive load, MANCOVA was employed to explore how various textual cues individually impacted two key cognitive dimensions—mental load and mental effort. The Bonferroni posthoc tests method was further used to scrutinize differences among groups. Additionally, ANOVA was leveraged to contrast and analyze the extent of change in students’ self-efficacy under diverse textual cues. Finally, correlation analyses were conducted by calculating correlation coefficients between post-test, delayed post-test, "faded effect" scores, and cognitive load and self-efficacy to elucidate the relationships among these variables.

4 Result

4.1 Learning outcomes

To answer the immediate and delayed effects of 3 types of textual cues on students’ learning outcomes after viewing educational animations. We conducted pre-tests, post-tests, and delayed post-tests from 3 groups. From the descriptive statistics, the pre-test scores of STC, KTC, and CTC were 2.20, 1.96, and 2.05, respectively, indicating that the students’ basic knowledge was weak. After having watched the educational animation, the post-test scores of the participants in the three groups

increased significantly, which were 7.03, 8.13, and 8.73, respectively. After 1 month, the delayed post-tests of the three groups decreased compared to the post-tests, which were 4.93, 6.29, and 7.36, respectively. The data for all groups passed Mauchly's test of sphericity ($W=0.995$, $\chi^2=1.367$, $p>0.05$). The results showed that the overall difference between the means of the three groups at different time points was significant ($F=47.420$, $p<0.001$, $\eta^2_p=0.291$), as shown in Fig. 3. This implies that there is a "faded effect" in the knowledge retention of the three groups. Therefore, the post-test scores, delayed post-test scores, and the "faded effect" of knowledge retention were analyzed separately for the three groups.

4.1.1 Analysis of the variability for achievement of three groups

To compare the achievement of the three groups, and to take into account the effect of prior knowledge, a one-way analysis of covariance was performed and all hypotheses were tested. The pre-test was used as the covariate, the group as the independent variable, and the post-test as the dependent variable ($F=31.584$, $p<0.001$, $\eta^2_p=0.201$). It can be seen that after controlling for the effect of the pre-test level, there is a significant difference in the amount of higher effect of achievement among the three groups.

To further investigate and statistically compare the post-test score differences among groups of participants, a covariance analysis was conducted in accordance

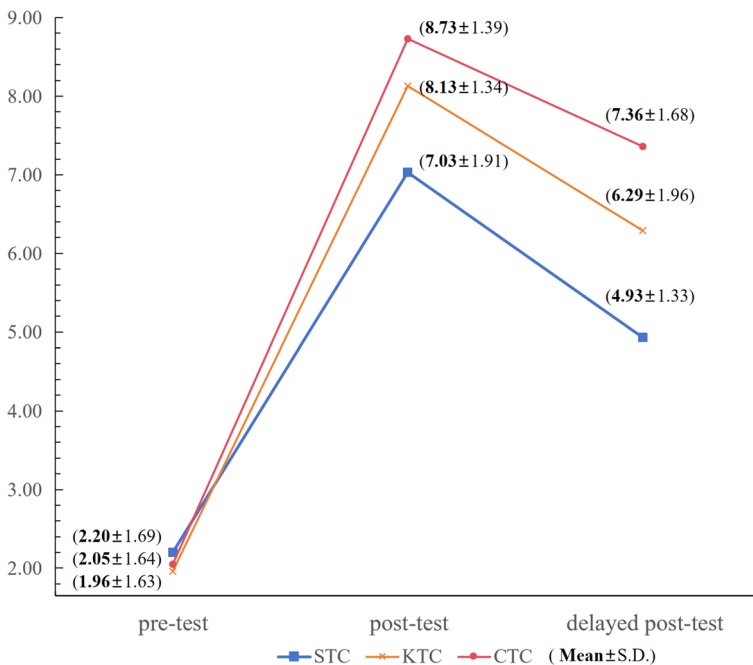


Fig. 3 Descriptive statistics and repeated measures analysis of variance for the pre-, post-, and delayed post-test of 3 groups

Table 2 The planning comparison result for post-tests of 3 groups

Group Comparison (I vs. J)	Adjusted Difference (I-J)	95%CI	<i>p</i>
STC vs. KTC	-1.11	(-1.69, -0.53)	.000***
KTC vs. CTC	-0.59	(-1.18, -0.01)	.000***
STC vs. CTC	-1.71	(-2.28, -1.13)	.000***

Covariate: pre-test; Dependent Variable: post-test

*** $p < .001$ **Table 3** Difference analysis results of knowledge retention of 3 groups

<i>F</i>	<i>p</i>	η^2_p	Group comparison (I vs. J)	Adjusted difference (I-J)	95%CI	<i>p</i>
46.459	.000***	0.268	STC vs. KTC	-1.46	(-2.10, -0.81)	.000***
			KTC vs. CTC	-1.09	(-1.75, -0.45)	.000***
			STC vs. CTC	-2.56	(-3.19, -1.92)	.000***

*** $p < .001$

with its methodological requisites (see Table 2). After controlling for the effects of the pre-tests, among the achievement of the 3 groups, the knowledge retention of the students in the CTC group were significantly higher than those of the KTC group ($p < 0.001$), and the knowledge retention of the KTC group were significantly higher than those of the STC group ($p < 0.001$). The above data indicates that the three groups have significant differences in the immediate test of learning outcomes and show the results of CTC > KTC > STC.

4.1.2 Analysis of variance in knowledge retention of three groups

A one-way ANOVA was used to analyze the delayed post-tests of the three groups, which showed a significant difference between the three groups with a large effect size ($F = 46.459$, $p < 0.001$, $\eta^2_p = 0.268$). To clarify the differences between groups, post hoc multiple comparisons were performed using the Bonferroni method. The results showed that students in the CTC group had significantly higher knowledge retention than those in the KTC group ($p < 0.001$), and the KTC group had significantly higher knowledge retention than the STC group ($p < 0.001$). The above data indicate that the three groups have significant differences in delayed post-tests and show the results of CTC > KTC > STC. The specific data are presented in Table 3.

4.1.3 Comparison of the "faded effect" of knowledge retention of three groups

To examine the distinct disparities in the "faded effect" of knowledge (KF) among the three groups, the current study computed each student's KF score (delayed post-test minus post-test) and executed a comparison. The KF score of every student was differentially calculated and compared in this investigation. Table 4 presents

Table 4 Comparison of "faded effect" of knowledge retention in 3 groups

KF	Mean \pm S.D	<i>F</i>	<i>p</i>	η^2_p	Group Comparison (I vs. J)	Adjusted Difference (I-J)	95%CI	<i>p</i>
STC (<i>n</i> = 87)	-2.11 \pm 1.96	7.931	.021*	0.061	STC vs. KTC	-0.36	(-1.14, -0.43)	.693
KTC (<i>n</i> = 84)	-1.85 \pm 2.02				KTC vs. CTC	-0.49	(-1.28, -0.53)	.382
CTC (<i>n</i> = 86)	-1.37 \pm 1.81				STC vs. CTC	-0.86	(-1.63, -0.08)	.024*

**p* < .05

the statistical depiction and analysis of the variance of KF among the three student groups. It demonstrated a significant difference in KF among the three groups with a medium effect size ($F=7.931$, $p<0.05$, $\eta^2_p=0.061$). Bonferroni posthoc comparisons method unveiled a statistically significant disparity solely between STC and CTC ($p<0.05$). There was no significant difference between KTC, and STC or CTC. However, it was evident that the |KF| of the 3 groups had the following order: STC ($M_{|KF|}=2.11$) > KTC ($M_{|KF|}=1.85$) > CTC ($M_{|KF|}=1.37$).

4.2 Cognitive load

MANCOVA was used to analyze the effects of applying STC, KTC and CTC in educational animation on students' cognitive load. Students' cognitive load was mainly reflected through mental load and mental effort. The descriptive statistics of the three groups of subjects are shown in Table 5.

Before analysis, the assumptions of the MANCOVA method were tested. The box plot test did not reveal any one-way outliers and the scatter plot test did not reveal any linear relationship between the dependent variables. The maximum Mahalano-bis distance is 13.519 (< 16.270, $\alpha=0.001$), indicating the absence of multivariate outliers. Box's M test showed $F=1.621$ ($p>0.001$), indicating equal covariance matrices for variance. As shown in the table, there was a mild correlation between the 2 variables of mental load and mental effort ($r=0.16$) and no multicollinearity (all $|r|<0.9$). The above results indicate that all hypotheses have been tested.

The results of the MANCOVA method of analysis are shown in Table 6. From a holistic perspective, the combined effect of textual cue type on students' cognitive load consisting of mental load and mental effort was significant ($F=21.357$, $p<0.001$, Wilks' $\lambda=0.652$; $\eta^2_p=0.233$), and more specifically, the one-way test

Table 5 Descriptive statistics of the cognitive load among the three groups

Group	Cognitive load	Mental load	Mental effort
	Mean \pm S.D	Mean \pm S.D	Mean \pm S.D
STC (<i>n</i> = 87)	3.67 \pm 0.962	3.52 \pm 1.082	3.92 \pm 1.333
KTC (<i>n</i> = 84)	3.02 \pm 1.152	2.86 \pm 1.132	3.29 \pm 1.012
CTC (<i>n</i> = 86)	2.58 \pm 0.998	2.43 \pm 0.979	2.83 \pm 0.819

Cognitive Load = 5/8 \times Mental Load + 3/8 \times Mental Effort

Table 6 Results of the MANCOVA for cognitive load: Multivariable test and Univariate tests

	<i>F</i>	<i>p</i>	η^2_p	Wilks' λ
Cognitive Load	21.357	.000***	0.233	0.652
Mental Load	17.577	.000***	0.122	
Mental Effort	16.564	.000***	0.115	

*** $p < .001$

analysis showed that there was a significant difference in the medium effect of mental load and mental effort for all three groups of subjects.

Pairwise comparisons were made using the Bonferroni posthoc tests method, and the results are shown in Table 7. For mental load, $STC > KTC > CTC$, and the results are significant. For mental effort, $STC > KTC > CTC$, and the results were significant. Cognitive load reflects the total amount of mental resources an individual requires to allocate and organize, representing the demands on the brain during the processing and retention of new knowledge (Paas & Merriënboer, 1994). Cognitive effort constitutes the actual portion of cognitive resources expended within the broader construct of cognitive load, denoting the actual intensity of cognitive processing invested by an individual during task execution. As the degree of task design rationality increases, the cognitive effort required from individuals diminishes correspondingly (Hwang et al., 2013). Consequently, in comparison with two other text-based cues, the presentation form in CTC appears relatively more rational and conducive to learners' comprehension of knowledge.

4.3 Self-efficacy

To investigate whether there exist significant differences in the effects of various types of educational animations on elementary school students' self-efficacy (SE), the ANOVA was employed (as presented in Table 8). The results revealed that indeed there are statistically significant differences in the influence of different types of educational animations on students' self-efficacy ($F = 23.423$, $p < 0.001$, $\eta^2_p = 0.141$).

Given this significant outcome, further investigation was conducted through the application of the Bonferroni post-hoc multiple comparisons test to ascertain

Table 7 Results of pairwise comparisons for cognitive load of 3 groups

Group Comparison(I vs. J)	Mental Load			Mental Effort		
	Adjusted Difference (I-J)	95%CI	<i>p</i>	Adjusted Difference (I-J)	95%CI	<i>p</i>
STC vs. KTC	0.66	(0.21, 1.08)	.001*	0.63	(0.15, 1.05)	.005*
STC vs. CTC	1.09	(0.64, 1.53)	.000***	1.09	(0.62, 1.53)	.000***
KTC vs. CTC	0.43	(0.01, 0.92)	.042*	0.46	(0.04, 0.95)	.036*

* $p < .05$, *** $p < .001$

Table 8 Results of ANOVA and Bonferroni post-hoc multiple comparisons for self-efficacy of 3 groups

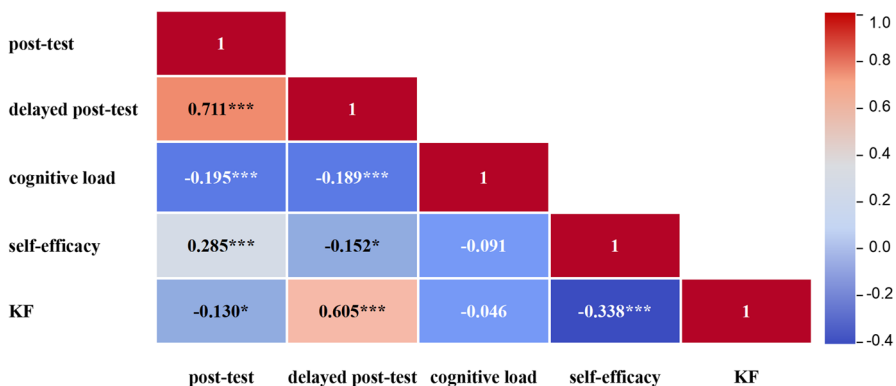
Descriptive statistics		ANOVA			Bonferroni post-hoc multiple comparisons			
Group	Mean \pm S.D	<i>F</i>	<i>p</i>	η^2_p	Group Comparison (I vs. J)	Adjusted Difference (I-J)	95%CI	<i>p</i>
STC	3.73 \pm 1.12	23.423	.000 ***	0.141	STC vs. KTC	-0.57	(-0.97,-0.17)	.003*
KTC	4.20 \pm 1.33				STC vs. CTC	-1.03	(-1.36,-0.69)	.000***
CTC	4.62 \pm 1.40				KTC vs. CTC	-0.46	(-0.86,-0.06)	.019*

* $p < .05$, *** $p < .001$

specific differences among the groups. Upon comparison, it was found that students exhibited the highest levels of self-efficacy when exposed to the educational animation with CTC, followed by the educational animation with CTC, while the educational animation with STC elicited the weakest self-efficacy scores, with all differences between these groups means being statistically significant.

4.4 Relationships among learning outcomes, cognitive load and self-efficacy

To investigate the relationship between learning outcomes, cognitive load and self-efficacy, we analyzed the correlations between post-test scores, KF, CL, and SE among three groups. Firstly, $r_{PT\&CL} = -0.195$ ($p < 0.001$) and $r_{DPT\&CL} = -0.189$ ($p < 0.001$), implying that the students' immediate post-test, delayed post-test all showed negative correlation with cognitive load, which verified the negative effect of cognitive load on the learning outcomes. $r_{KF\&SE} = -0.338$ ($p < 0.001$), This indicates that the results of the delayed post-test are negatively correlated with self-efficacy. in addition, $r_{DPT\&SE} = -0.152$ ($p < 0.05$), indicates that the "faded effect" related



Note: * $p < .05$, *** $p < .001$; PT: post-test, DPT: delayed post-test, CL: cognitive load, SE: self-efficacy, KF: the "faded effect" of knowledge retention.

Fig. 4 Correlation analysis of learning outcomes, cognitive load and self-efficacy

to knowledge retention has a significant negative correlation with students' self-efficacy. Since the mean self-efficacy of the three groups in this experiment was high after viewing the educational animation (mean = 4.18), excessive self-efficacy on the part of the students may hinder the knowledge retention. Detailed data are presented in Fig. 4.

5 Discussion

Educational animation engages students and promotes understanding of complex concepts by providing visual and auditory stimulation. However, students' learning outcomes and experience using these educational animations are affected by the type of textual cues that accompany them. This study investigated the impact of different textual cues—subtitle textual cue (STC), keyword textual cue (KTC), and structured textual cue (CTC)—on student achievement, knowledge retention, cognitive load, and self-efficacy.

In terms of achievement, students using educational animations with CTC achieved the highest scores, followed by those using KTC, with STC performing the worst. The poor academic achievement associated with STC could be attributed to visual distraction caused by dynamic subtitles, which obscured key information (Arslan-Ari et al., 2020; De Koning & Jarodzka, 2017). Additionally, based on the transient information effect (Castro-Alonso et al., 2018), the fleeting nature of subtitles can lead to cognitive overload, impeding students' ability to process and retain information effectively. In contrast, students using CTC in educational animations outperformed those using KTC. This supports previous research that constructing knowledge from animations is an iterative cognitive process (Kriz & Hegarty, 2007), and well-structured textual cue can more clearly guide this process. CTC, with its hierarchical structure, facilitates the knowledge development process, leading to meaningful learning (Semilarski et al., 2022).

Regarding knowledge retention, all groups showed lower delayed post-test scores compared to immediate post-test scores, indicating the "faded effect". While the CTC group scored slightly higher than the KTC group, the difference was not significant, and the STC group lagged significantly behind the other two. This interesting finding suggests that the positive impact of educational animations on students may diminish over time, regardless of the type of textual cues used. This aligns with Ebbinghaus's forgetting curve (Ebbinghaus, 2013). The significantly lower performance of the STC may be due to the transient information effect (Castro-Alonso et al., 2018), where the fleeting nature of subtitles leads to cognitive overload, hindering students' ability to process and retain information effectively. The dynamic visual nature of STC introduces additional cognitive burdens, making comprehension more difficult over time, thus resulting in lower retention rates (Paas et al., 2007).

Regarding self-efficacy, students using CTC reported the highest self-efficacy scores, followed by those using KTC, with STC yielding the lowest self-efficacy scores. For students, constructing knowledge from animations is essentially an iterative cognitive process (Kriz & Hegarty, 2007). CTC could more clearly use textual cues to organize the semantic content and structure of the material, guiding students'

cognitive processes and enhancing their confidence in their learning abilities (Sweller, 2010). While KTC had a lower degree of structuring, its clear keywords still enabled students to immediately notice the key points of the course, thereby enhancing self-efficacy (Semilarski et al., 2022). Students using STC reported the lowest self-efficacy, possibly due to cognitive overload caused by subtitles, which hindered their engagement and confidence in the learning tasks (Paas et al., 2007).

In terms of the correlations between learning outcomes, cognitive load, and self-efficacy, the study found a significant negative correlation between cognitive load and both academic achievement and knowledge retention, consistent with previous research (Arslan-Ari et al., 2020; De Koning & Jarodzka, 2017; Semilarski et al., 2022). However, we also found an interesting correlation regarding self-efficacy. Students' self-efficacy was significantly positively correlated with their learning outcomes, but it also showed a significant correlation with the faded effect. Through discussions with students and their science teachers, it became evident that many students in the CTC group believed they had thoroughly grasped the concept of Earth's rotation after watching the educational animation. Consequently, they did not feel the need for further review. When students perceive themselves as having rich learning experiences in a specific medium, feel highly confident about their learning situation, and optimistically believe they can acquire extensive knowledge in some way, they may misjudge the difficulty of the task (Toni et al., 2023) or erroneously believe they do not need further guidance or review (Acuna et al., 2011). We observed that students in all groups reported high self-efficacy in their learning experiences. In traditional science classrooms, students' self-efficacy is considered a crucial indicator of teaching effectiveness (Lin, 2021; Tarchi et al., 2021). When students have high self-efficacy, they are generally expected to exert more effort in science learning and have higher cognitive pursuits (Semilarski et al., 2022). However, this study's findings extend previous research on science learning, suggesting that, for indigenous digital students, self-efficacy may be a cautionary indicator in measuring multimedia-supported science education. Teachers should provide more metacognitive support and materials for review and reflection to address this potential challenge.

6 Implications

The findings underscore the critical role of textual cue design in educational animations. Effective cue design can enhance learning outcomes, improve knowledge retention, manage cognitive load, and boost self-efficacy, providing valuable insights for educators and instructional designers in creating optimized multimedia learning environments. Meanwhile, the study highlights the need for a nuanced approach to leverage self-efficacy for multimedia-assisted science education, emphasizing the importance of designing instructional materials that simultaneously leverage students' confidence while addressing the potential pitfalls of overinflated self-efficacy in sustaining long-term learning outcomes. To this end, the present study proposes the following recommendations for the design of educational animation:

- Educational animation should employ textual cues that possess distinct semantics and a well-structured format, such as CTC, to facilitate students' cognitive development and the establishment of cognitive schema.
- The educational effectiveness of transient information, such as dynamic subtitles (STC), in educational animation is less than satisfactory. Therefore, it is advised to minimize its usage. However, if its utilization is deemed necessary, it is vital to adopt more nuanced instructional designs that enable students to concentrate on pertinent information.
- When managing cognitive load, educators can consider highlighting key information by marking keywords (such as KTC) and placing textual cues in fixed positions that align with reading habits (such as CTC). These methods help reduce cognitive load and improve learning effectiveness.
- When digital native students engage in multimedia with a lower cognitive load, such as CTC educational animation, they are often inclined to exhibit a heightened sense of self-efficacy, which may result in overconfidence in the learning outcomes and a disregard for subsequent review and consolidation. Thus, it is crucial for teachers to remind and guide students toward conducting effective reviews and knowledge retention. Alternatively, teachers can enhance students' metacognition by incorporating additional cues in instructional design or educational animation.

7 Conclusion

This study highlights the significant impact of textual cues on students' learning outcomes, cognitive load, and self-efficacy in educational animations. The findings suggest that structured textual cues (CTC) are most effective in enhancing student achievement and self-efficacy while minimizing cognitive load. Keyword textual cues (KTC) also provide benefits, though to a lesser extent, while subtitle textual cues (STC) may introduce additional cognitive challenges that hinder learning. Additionally, the study underscores the complex role of self-efficacy in learning, indicating that high self-efficacy may sometimes lead to overconfidence and decreased review efforts, potentially impacting knowledge retention. These insights provide valuable guidance for educators and instructional designers in creating effective educational animations that support meaningful learning and sustained student engagement.

8 Limitations and future

Although significant effects were achieved, this study still has some limitations. Firstly, the educational animations used in this research were presented to students in a systematically controlled manner, depriving them of control over the pace of the animations. Previous studies (Tabbers & De Koeijer, 2010) have suggested that allowing students to control their learning pace can lead to better learning performance. Future investigations could explore the impact of students autonomously

controlling educational animations. Secondly, this study investigated how three different types of textual cues impact various learning outcomes concerning self-efficacy and cognitive load from the perspective of the learner's learning experience. Acknowledging the critical role that text quantity plays in achieving efficacious educational outcomes through animated instruction (Berney & Betrancourt, 2016), we advocate for further research into designing an appropriate amount of text for each cue type, aiming to ensure that essential information is conveyed effectively, thereby optimizing learners' usage experience and learning effectiveness. Thirdly, we suggest that future researchers extend the duration of the experimental period and develop a series of educational animations with the three types of textual cues to observe the long-term effects on students' learning outcomes and learning experiences.

Funding “Big Data Education Evaluation Research”, the priority topic of the 14th Five-Year Plan for Education Science of Beijing, China in 2022 (No.CDEA22008).

Data availability The datasets used and analyzed in the current study are available on reasonable request.

Declarations

Ethical approval All procedures performed in studies involving human participants followed the ethical standards of the institutional and national research committees. Participation in this study was voluntary. All participants were informed about the study and signed a free informed consent.

Conflict of Interest There is no potential conflict of interest in this study.

Competing interests The authors declare that they have no competing interests.

References

- Acuna, S. R., Rodicio, H. G., & Sanchez, E. (2011). Fostering active processing of instructional explanations of learners with high and low prior knowledge. *European Journal of Psychology of Education*, 26(4), 435–452. <https://doi.org/10.1007/s10212-010-0049-y>
- Alpizar, D., Adesope, O. O., & Wong, R. M. (2020). A meta-analysis of signaling principle in multimedia learning environments. *Educational Technology Research and Development*, 1–25. <https://doi.org/10.1007/s11423-020-09748-7>
- Arslan-Ari, I., Crooks, S. M., & Ari, F. (2020). How much cueing is needed in instructional animations? The role of prior knowledge. *Journal of Science Education and Technology*, 29(5), 666–676. <https://doi.org/10.1007/s10956-020-09845-5>
- Barak, M., Ashkar, T., & Dori, Y. J. (2011). Learning science via animated movies: Its effect on students' thinking and motivation. *Computers & Education*, 56(3), 839–846. <https://doi.org/10.1016/j.compedu.2010.10.025>
- Berney, S., & Betrancourt, M. (2016). Does animation enhance learning? A meta-analysis. *Computers & Education*, 101, 150–167. <https://doi.org/10.1016/j.compedu.2016.06.005>
- Castro-Alonso, et al. (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>
- Castro-Alonso, J. C., Ayres, P., Wong, M., & Paas, F. (2018). Learning symbols from permanent and transient visual presentations: Don't overplay the hand. *Computers & Education*, 116, 1–13. <https://doi.org/10.1016/j.compedu.2017.08.011>

- Clark, R. C., & Mayer, R. E. (2016). *e-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning*. John Wiley & Sons. <https://doi.org/10.1002/9781119239086>
- De Koning, B. B., & Jarodzka, H. (2017). Attention guidance strategies for supporting learning from dynamic visualizations. In Richard Lowe & Rolf Ploetzner (Eds.), *Learning from dynamic visualization* (pp. 255–278). Springer. https://doi.org/10.1007/978-3-319-56204-9_11
- 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>
- Dunning, D. (2011). The Dunning-Kruger effect: On being ignorant of one's own ignorance. *Advances in Experimental Social Psychology*, 44, 247–296. <https://doi.org/10.1016/B978-0-12-385522-0.00005-6>
- Ebbinghaus, H. (2013). Memory: A contribution to experimental psychology. *Annals of Neurosciences*, 20(4), 155–156. <https://doi.org/10.5214/ans.0972.7531.200408>
- Granic, A. (2022). Educational technology adoption: A systematic review. *Education and Information Technologies*, 27(7), 9725–9744. <https://doi.org/10.1007/s10639-022-10951-7>
- Hwang, G. J., Yang, L. H., & Wang, S. Y. (2013). A concept map-embedded educational computer game for improving students' learning performance in natural science courses. *Computers & Education*, 69, 121–130. <https://doi.org/10.1016/j.compedu.2013.07.008>
- Karlsson, G. (2010). Animation and grammar in science education: Learners' construal of animated educational software. *International Journal of Computer Supported Collaborative Learning*, 5(2), 167–189. <https://doi.org/10.1007/s11412-010-9085-5>
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.
- Kintsch, W. (2019). Revisiting the construction–integration model of text comprehension and its implications for instruction. In D. E. Alvermann, N. J. Unrau, & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (7th ed., pp. 178–203). International Reading Association. <https://doi.org/10.4324/9781315110592-12>
- Kriegelstein, F., Schneider, S., Gröninger, J., Beege, M., Nebel, S., Wesenberg, L., Suren, M., & Rey, G. D. (2023). Exploring the effects of content-related segmentations and metacognitive cues on learning with whiteboard animations. *Computers & Education*, 194, <https://doi.org/10.1016/j.compedu.2022.104702>
- Kriz, S., & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human-Computer Studies*, 65(11), 911–930. <https://doi.org/10.1016/j.ijhcs.2007.06.005>
- Lin, T. J. (2021). Multi-dimensional explorations into the relationships between high school students' science learning self-efficacy and engagement. *International Journal of Science Education*, 43(8), 1193–1207. <https://doi.org/10.1080/09500693.2021.1904523>
- Lowe, R. K., & Schnotz, W. (2014). Animation principles in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 513–546). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.026>
- Matthew, G. (2020). The effect of adding same-language subtitles to recorded lectures for non-native, English speakers in e-learning environments. *Research in Learning Technology*, 28, 2340–2355. <https://doi.org/10.25304/rlt.v28.2340>
- Mayer, R. E. (2024). The past, present, and future of the cognitive theory of multimedia learning. *Educational Psychology Review*, 36, 8. <https://doi.org/10.1007/s10648-023-09842-1>
- Mayer, R. E., & Pilegard, C. (2014). Principles for managing essential processing in multimedia learning: Segmenting, pre-training, and modality principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 316–344). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.016>
- Ministry of Education (MOE). (2022). *Science curriculum standards for compulsory education*. Beijing Normal University Press. (In Chinese).
- Nguyen, H. H., Do Trung, K., Duc, L. N., et al. (2024). A model to create a personalized online course based on the student's learning styles. *Education and Information Technologies*, 29, 571–593. <https://doi.org/10.1007/s10639-023-12287-2>
- Paas, F., Van Gerven, P. W. M., & Wouters, P. (2007). Instructional efficiency of animation: Effects of interactivity through mental reconstruction of static key frames. *Applied Cognitive Psychology*, 21, 783–793. <https://doi.org/10.1002/acp.1349>

- Paas, F. G. W. C., & Van Merriënboer, J. J. G. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review*, 6(4), 351–371. <https://doi.org/10.1007/BF02213420>
- Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies questionnaire (MSLQ)*. <https://doi.org/10.13140/RG.2.1.2547.6968>
- Ploetzner, R., Berney, S., & Bétrancourt, M. (2021). When learning from animations is more successful than learning from static pictures: Learning the specifics of change. *Instructional Science*, 49, 497–514. <https://doi.org/10.1007/s11251-021-09541-w>
- Ploetzner, R., & Lowe, R. (2012). A systematic characterisation of expository animations. *Computers in Human Behavior*, 28(3), 781–794. <https://doi.org/10.1016/j.chb.2011.12.001>
- Richter, J., Scheiter, K., & Eitel, A. (2018). Signaling text–picture relations in multimedia learning: The influence of prior knowledge. *Journal of Educational Psychology*, 110(4), 544–560. <https://doi.org/10.1037/edu0000220>
- Rop Schüller, A., Verkoeijen, P. P., Scheiter, K., & Gog, T. V. (2018). The effect of layout and pacing on learning from diagrams with unnecessary text. *Applied Cognitive Psychology*, 32(5), 610–621. <https://doi.org/10.1002/acp.3445>
- Semilarski, H., Soobard, R., Holbrook, J., & Rannikmäe, M. (2022). Expanding disciplinary and interdisciplinary core idea maps by students to promote perceived self-efficacy in learning science. *International Journal of STEM Education*, 9(1), 1–20. <https://doi.org/10.1186/s40594-022-00374-8>
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, 22(2), 123–138. <https://doi.org/10.1007/s10648-010-9128-5>
- Tabbers, H. K., & De Koeijer, B. (2010). Learner control in animated multimedia instructions. *Instructional Science*, 38, 441–453. <https://doi.org/10.1007/s11251-009-9119-4>
- Tarchi, C., Zaccoletti, S., & Mason, L. (2021). Learning from text, video, or subtitles: A comparative analysis. *Computers and Education*, 160, <https://doi.org/10.1016/j.compedu.2020.104034>
- Toni, H., Jaclyn, B., & Matthew, F. (2023). The self-efficacy and academic performance reciprocal relationship: The influence of task difficulty and baseline achievement on learner trajectory. *Higher Education Research & Development*, 42(8), 1936–1953. <https://doi.org/10.1080/07294360.2023.2197194>
- Tosun, C. (2022). Analysis of the last 40 years of science education research via bibliometric methods. *Science and Education*, 1–30. <https://doi.org/10.1007/s11191-022-00400-9>
- Türkyay, S. (2016). The effects of whiteboard animations on retention and subjective experiences when learning advanced physics topics. *Computers & Education*, 98, 102–114. <https://doi.org/10.1016/j.compedu.2016.03.004>
- Wang, F., Zhao, T., Mayer, R. E., & Wang, Y. (2020). Guiding the learner’s cognitive processing of a narrated animation. *Learning and Instruction*, 69, 1–12. <https://doi.org/10.1016/j.learninstruc.2020.101357>
- Yilmaz. (2023). Effects of using cueing in instructional animations on learning and cognitive load level of elementary students in science education. *Interactive Learning Environments*, 31(3), 1727–1741. <https://doi.org/10.1080/10494820.2020.1857784>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.