Education and Information Technologies (2024) 29:16929–16956
https://doi.org/10.1007/s10639-024-12540-2

# Mental imagery scaffolding: The effects of detail richness and text load on geography learning

**Yun Zhou[1] · Fanqi Yi[1] · Bingyu Dong[1] · Guangli Zhang[1] · Yi Zhang[2] · Tao Xu[2]**

Received: 4 July 2023 / Accepted: 5 February 2024 / Published online: 17 February 2024


## Abstract
The growing importance of 3D animations in current teaching approaches becomes increasingly apparent, offering an effective way to visualize complex spatial concepts and processes in geography learning through outstanding visual representation and details. However, the effects of detail richness and text load of 3D animation on learning about processes remain unclear. Addressing this research gap, the present study adopts a quasi-experimental design involving four classes ($n=106$) in the context of a geography lesson and evaluates four conditions in a $2\times2$ between-subjects design consisting of detail richness (high vs. low) and text load (high vs. low). The lessons on the rotation and revolution of the Earth were delivered by the same instructor across all conditions. Knowledge acquisition, cognitive load, learning experience, and emotions of students were measured. The results revealed that students were significantly better able to acquire knowledge immediately when exposed to the high detailed visuals but low text load condition. Low detail richness and high text load independently resulted in increased cognitive load. We also observed a significant effect of detail richness on the dimensions of pleasure and arousal, with higher levels of details associated with larger values in these dimensions. This research suggests that when the learning objective necessitates the engagement of mental imagery, incorporating detailed visuals can facilitate learning. The findings contribute to our understanding of how detailed imagery is linked to learning objectives about processes and expand our knowledge regarding the design of detail richness and text of 3D animation in the context of geography learning.

**Keywords** 3D animation · Detail richness · Visual representation · Realism · On-screen text · Geography learning

✉ Tao Xu
xutao@nwpu.edu.cn

[1] Faculty of Education, Shaanxi Normal University, South Chang'an Road 199, Yanta District, Xi'an, Shaanxi Province 710062, P.R. China

[2] School of Software, Northwestern Polytechnical University, 127 West Youyi Road, Beilin District, Xi'an, Shaanxi Province 710072, P.R. China

## 1 Introduction

Visual representations are essential for improving STEM (science, technology, engineering, and mathematics) learning (Rau, 2017). Within the STEM disciplines, learning often centers around the comprehension of complex processes. For instance, in geography, students engage in the study of phenomena like the Earth's rotation and revolution. To facilitate comprehension of such concepts, the transition from traditional 2D learning resources to the 3D animations emerges as a transformative pedagogical approach (Wu & Chiang, 2013). These 3D animations offer students a novel window into the dynamic processes or systems, visualizing the complex spatial information that were previously challenging to convey effectively. As educational paradigms evolve, the growing importance of 3D animations in current teaching approaches becomes increasingly apparent. These animations leverage the perceptual richness intrinsic to their design, transforming themselves into effective tools for generating enhanced geography learning experiences. However, even as the educational landscape embraces the potential of 3D animations, certain questions remain unexplored within the existing literature. The effect pattern of realism of visual representations when learning tasks involving understanding processes or systems is largely unknown (Skulmowski & Rey, 2021a). Additionally, most studies concentrate on examining the effects of learning performance (e.g., Scheiter et al., 2009; Menendez et al., 2020) and cognitive load (Birbara & Pather, 2021; Joo et al., 2021; Skulmowski, 2022c), leaving the impact of incorporating detailed visuals on the emotional aspects of learning unclear.

In addition to visual features, on-screen text in 3D animations is useful for explaining complicated geographical phenomena and assisting students in understanding crucial concepts. Previous research indicates that the integration of text into animation can enhance students' learning when compared to text-only version (e.g., Kablan & Erden, 2008). The incorporation of text alongside visual elements in 3D animation offers a dual channel for conveying information, capitalizing on both the textual and visual modes of comprehension (Moreno & Mayer, 1999b, 2002; Butcher, 2006). However, despite the advantages of combining text and visuals in 3D animations, several questions remain unanswered. First, the optimal amount of text, referred to as the "text load," within the context of 3D animations is still being researched. It's essential to determine the balance between text and visual elements to ensure that the information presented stays comprehensible without overwhelming the learner. Furthermore, the interaction between the different levels of detail richness in 3D animations and the text load requires investigation. By exploring the interplay between text load and detailed imagery within 3D animations, this study aims to shed light on the most effective strategies for presenting information in a way that optimizes the learning experience in geography course.

Building upon this background, this study examines the effects of detail richness and text load of 3D animation on learning about process in the context of geography learning.

### 1.1 Research on the effects of visual representation and details

In multimedia learning, visual representations, such as diagrams, charts, graphs, illustrations, videos, or animations, are used to represent another objects, ideas, or concepts by incorporating similar features to the original entities (Rau, 2017, 2020). With the advancements in computer graphics and multimedia technology, researchers have increasingly recognized the impact of visual realism and the richness of representation details on the learning process. This recognition has given rise to two contrasting perspectives on the potential benefits of visual realism and rich detail for learning. On one hand, some researchers believe that realistic visualization may have a negative impact on learning due to its high demands on visual attention. For example, findings from Experiment 1 by Scheiter et al. revealed that learners felt overwhelmed by the excessive amount of realistic details, hindering their ability to accurately comprehend the process of mitosis. In comparison, students exposed to schematic visualizations outperformed those exposed to realistic representations (Scheiter et al., 2009), as schematic visualizations focus on primary features and facilitate interpretation of represented objects. Naive realism also suggests the use of simplified imagery over realistic representations (Smallman & John, 2005). According to cognitive load theory (Sweller et al., 2019), which has been regarded as a theoretical foundation guiding visual representations research, unnecessary and complicated details may result in increased extraneous cognitive load (Castro-Alonso et al., 2019). One objective of multimedia instructional design principles is to reduce extraneous cognitive load. Consequently, the inclusion of realistic details is not recommended (Belenky & Schalk, 2014). Furthermore, some empirical studies have found that realistic visualizations have no impact on learning. Both experts and novices may be drawn to rich visualizations, but their performance did not meet expectations (Hegarty et al., 2012). Even though rich detail was relevant to the task, students learning about biological change with the rich diagram did not outperform those using the abstract diagram in terms of transfer performance (Menendez et al., 2020). Although physical fidelity has the potential to improve knowledge outputs, it is not necessarily required in virtual learning materials (Birbara & Pather, 2021).

On the other hand, recent studies have argued that realistic details can positively influence learning. Researchers who support this perspective have found that, in learning scenarios requiring interaction with real objects in a natural setting, realistic visual representations effectively engage students and hold their appeal. Harrington found that using highly realistic virtual environments combined with navigational freedom has a positively significant impact on learning outcomes and can stimulate greater curiosity (Harrington, 2012). Waller et al. also found that a detailed virtual environment (VE) helps subjects transfer their spatial knowledge more effectively compared to an undetailed VE (Waller et al., 2001). Additionally, researchers have found that high realism has a positive impact on memory tests and image assessments during learning, even though it may increase cognitive load. Using more realistic visualizations can effectively improve retention in learning tasks that emphasize the memorization of shapes and are evaluated through image-based tests (Skulmowski, 2022b). Realistic visualizations have also been linked to enhanced performance in transfer tasks (Skulmowski, 2022b).

The aforementioned findings reveal that the impact of visual realism remains unpredictable (Skulmowski, 2023), and there is a lack of comprehensive research on its pattern. While realism may aid surface-related knowledge, it could hinder abstract understanding (Skulmowski & Rey, 2021a), and its effect on tasks related to understanding processes remains unclear. Many STEM subjects such as geography require an understanding of processes or systems. Therefore, further investigation is required to better understand the effect patterns of detail richness and whether the realistic dynamic representation facilitates the acquisition of conceptual knowledge about processes or systems.

Evaluating learning experience and emotions alongside learning performance is essential for a comprehensive understanding of the learning. Learning experience involves the quality of engagement, motivation, and interest that learners experience throughout the learning process (Awidi & Paynter, 2019). By assessing learning experience, educators and stakeholders can identify the factors that enhance or hinder student engagement, allowing them to make necessary adjustments to optimize. For learning media or environments that incorporate dynamic visualization, such as Virtual Reality (VR), evaluating the learning experience holds significant importance and has gained widespread attention (Ai-Lim Lee et al., 2010; Guerra-Tamez, 2023). The effectiveness of these dynamic visualizations relies heavily on how learners perceive and engage with them (Di Natale et al., 2020). In the learning process, cognitive and emotional processes intricately interact, shaping the overall learning experience and influencing learning outcomes (Pekrun & Linnenbrink-Garcia, 2012). Emotions have a profound impact on various stages of learning, from attention and encoding to information processing and retrieval (Han et al., 2021). Recent studies concentrate on examining the effects of learning performance (e.g., Birbara & Pather, 2021; Skulmowski, 2022a) and cognitive load (e.g., Skulmowski, 2023). While the credibility of visualizations (Skulmowski & Rey, 2021b) has been addressed, the influence of detail richness on the experiential and emotional aspects of learning remains unexplored. To acquire a deeper understanding of the effects of detailed imagery, we therefore investigated learners' learning experience and emotions in this study, in addition to knowledge acquisition and cognitive load.

### 1.2 Research on the integration of text and visual information

Understanding complex 3D animations can be challenging in the absence of appropriate on-screen text, which, in multimedia learning, encompasses any textual content displayed on a screen. On-screen text exists in various forms such as subtitles, headings, prompts, and annotations. They serve multiple functions, such as offering translations (Tarchi et al., 2021), explanations (Plass et al., 1998), descriptions (Li et al., 2019), and textual navigation guidance within interactive interfaces (Chen, 2020). Multimedia learning resources may also include explained content to provide detailed information, explanations, or examples related to the topic being presented. These explanations are often used to convey complex information, concepts, or details in a more structured way (Ardaç & Unal, 2008). In multimedia learning, on-screen text plays a crucial role in enhancing learner engagement, knowledge acquisition, and comprehension of information.

When animations are employed, visual forms are not the only options to stimulate deep levels of knowledge processing among learners. The integration of text into animation has been focused and explored. Previous studies have shown that students learned with integrated text and animation format performed better than students in the group with separated presentation format (e.g., Kablan & Erden, 2008). According to the multimedia principle, the most effective way to learn diagrams or animations is by presenting them in conjunction with text (Lowe & Schnotz, 2014). The integration of text and graphics stimulates a deeper level of knowledge organization processing among learners than using words alone (Mayer & Anderson, 1991, 1992; Mayer, 2003). When text and graphics are given together, students can build verbal and graphical mental models and make connections between them. However, the redundancy principle suggests using either speech or text when explaining diagrams, but not both simultaneously. While the redundancy principle argues that on-screen text may hinder learning, ongoing research continues to explore its boundaries. Some studies have found that brief text prompts, text annotations, or summary texts can actually enhance learning (Mayer & Johnson, 2008; Yue et al., 2013). Conversely, it has been observed that the full-text version tends to benefit students more (Ardaç & Unal, 2008). The optimal amount of text for effective learning to be presented alongside graphics and animations remains unclear.

In geography learning, where processes and spatial relationships are central, visual aids in 3D animations become important. Besides visual details, on-screen text in these animations also serves as a valuable tool to assist in clarifying complex geographical phenomena and aiding students in comprehending key concepts. However, the interaction between the different rich levels of detailed imagery and the text load in 3D animation is unknown according to the existing literature. Does a highly detailed 3D animation benefit from more text-based explanations, or does it require less text due to visual richness? On the other hand, in undetailed animations, does text play a more prominent role in conveying information? Taken together, this study examined the impact of text load on learning and explored the interaction between detail richness and text load.

### 1.3 3D animation is an effective tool for facilitating geography learning

The rapid development of information technology has made 3D animation an effective tool for facilitating learning (Lowe & Schnotz, 2014; Berney & Bétrancourt, 2016). Its diverse applications encompass several domains, such as architecture, physics, chemistry, electronics, anatomy, astronomy, and biology. Empirical studies have demonstrated that the use of 3D animation facilitates the visualization of complex concepts and processes, significantly improving students' knowledge acquisition and understanding of the material (Ploetzner et al., 2020).

In addition, 3D animated learning resources can serve an interpretive function in revealing invisible relationships, showing an object or phenomenon from different or changing perspectives (Hegarty, 2004). The visual appeal and interactive nature of these resources are effective in capturing students' attention and maintaining their engagement throughout the lesson. The unique stereoscopic vision and dynamic features of 3D animation provide a superior visual experience compared to conventional 2D learning resources (Wu & Chiang, 2013).

Given the challenges in learning geography, including grasping abstract concepts, and understanding spatial relationships and the complexity of geographic processes at different scales (Rickey & Bein, 1996; El-Nahass & Abdellatif, 2021), the aforementioned benefits of 3D animation become particularly valuable in geography learning. Building upon this background, our study seeks to investigate the effects of detail richness and text load of 3D animations on the comprehension of geographical processes. Our objective is to identify the most effective strategies for presenting information that enhances the overall learning experience in geography courses.

### 1.4 The terminology used in this study

In this paper, the term "detail richness" refers to the visual realism of visual representation, animated effects, and visual cues. Visual representation involves the various elements and characteristics contributing to the overall appearance and realism of the rendered animation. Animated effects refer to the process of creating movement and changing the appearance of objects within an application. Visual cues are the use of visual aids to improve the learning process, such as color shifts and arrows. These details are created using digital tools and techniques to simulate the appearance and motions of real-world objects, environments, and characters, as well as to direct attention and demonstrate relations.

The term "text load" refers to the amount of text or written content displayed on a screen. It includes three aspects: quantity, complexity, and type of text. The design of low/high detail richness and low/high text load will be depicted respectively in subsections of the 3D animation resources.

### 1.5 Present study and research questions

This study aims to fill the aforementioned gap by examining the effects of detail richness and text load on performance, cognitive load, learning experience, and emotions in geography learning.

The study will explore the following four research questions:

RQ1: How do detail richness (low/high) and text load (low/high) affect the acquisition of knowledge?

RQ2: How do detail richness and text load affect the cognitive load?

RQ3: How do detail richness and text load affect the learning experience and emotion?

RQ4: Do detail richness and text load interact in their impact on learning?

## 2 Methods

### 2.1 Participants and research design

This work was conducted using a quasi-experimental research method. Four classes of seventh-grade students, comprising a total of 106 students (53 boys and 53 girls), participated in the study. They were aged between 12 and 13. This study adopts a 2×2 between subjects design. The independent variables manipulated were detail richness of 3D animation (high/low) and text load (high/low). Four classes were randomly assigned to four experimental conditions, which were: HDHT (high detail richness with high text load, $n=26$), HDLT (high detail richness with low text load, $n=27$), LDHT (low detail richness with high text load, $n=26$), and LDLT (low detail richness with low text load, $n=27$). All participants had normal or corrected to normal visions. All experimental groups were exposed to the same lesson with only the details of 3D animated resources and the amount of text varied. The lessons on the rotation and revolution of the Earth were delivered by the same instructor across all conditions. This study was approved by the ethics board of our institution. The school and students' parents were informed about this investigation and asked for an informed consent agreement.

A power analysis was conducted using G*Power 3.1 (Faul et al., 2007) based on the effect size (0.4) found in (Kang, 2021). With an effect size of f=0.4, power of 0.80, and alpha level of 0.05, a total 80 participants were needed for four groups to detect an effect (i.e., 20 for each group). Therefore, the number of participants for each group in this study met the requirements.

### 2.2 The 3D animation resources

The 3D animated instructional resources used in this study were developed using 3Ds Max. The subject of the course was the rotation and revolution of the Earth. The 3D animated resources covered three important aspects: (1) the location and motion of the Earth, Sun, and other planets within the solar system; (2) the demonstration of the Earth's rotation from multiple perspectives; and (3) the Earth's revolution and its four positions on the orbital plane.

#### 2.2.1 Detail richness

All experimental groups were exposed to the same lesson, with variations only in the details of the 3D animated resources and the amount of text. In this study, the detail richness was designed and developed based on the model shown in Fig. 1, which includes objects, granularity aspects, animation, visual cues, and resolution. To avoid ambiguity, we have specified these elements.

Everything in the 3D world is considered an object and can be classified as a character, item, or environment. A character is typically a 3D model, humanoid or non-humanoid (like robots), representing a player or a non-player character. Characters can be controlled by the player or the computer, interacting with other objects in the 3D world. An item is any manipulable element, such as a cube. The environment

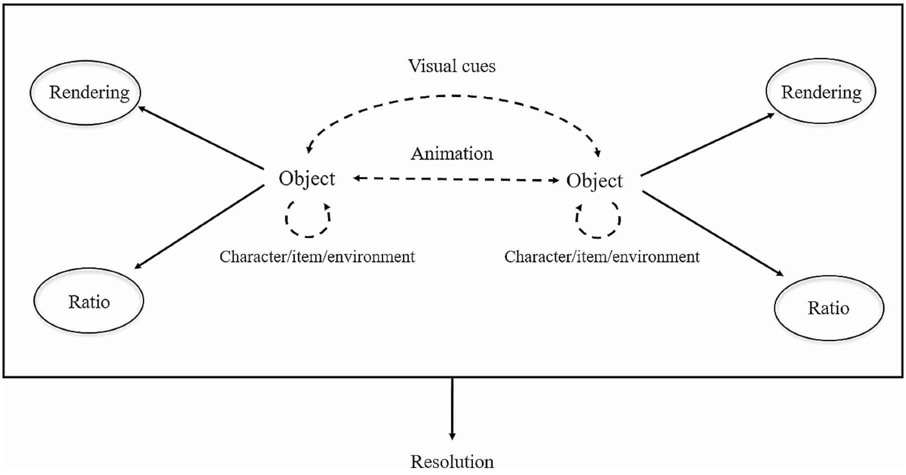


**Fig. 1** A model depicting aspects considered in detail richness

refers to the 3D world or scene created to house and contain the developed application being developed. The environment refers to the 3D world or scene created to house and contain the developed application, comprising various elements like terrain, objects, buildings, characters, and other visual and interactive components.

Granularity aspects of rendering and ratio are used to illustrate the detail richness of each object. Rendering refers to the process of generating images from 3D models and scenes, involving the use of a graphics engine that combins various elements such as textures, shaders, and geometry. The ratio of objects in 3D world refers to the size and scale of virtual objects in relation to the user's perception of real-world objects.

Animated effects refer to the process of creating movement and changing the appearance of objects within an application. They can indicate static or dynamic relationships between objects.

Visual cues are aids used in educational settings to enhance the learning experience and improve understanding of concepts, often in the form of drawings or texts. They can indicate static or dynamic relationships between objects.

Resolution refers to the number of pixels or dots that make up an image on a display device such as a computer monitor. A higher resolution means that more pixels are used to create the image, resulting in a sharper and more detailed image with smoother lines and curves.

On the one hand, the 3D animations with low-level detailed imagery feature a plain colored background and use images with a simple texture to render celestial bodies such as the Earth. These animations are exported at lower resolutions and include only a limited number of visual cues. There are also fewer effects to convey dynamic changes in the animation, and the level of detail is minimal. Only essential details within the animations were presented.

On the other hand, in 3D animations with a high level of visual detail, the background features a dynamic starry sky, adding a sense of realism. Celestial bodies, like the Earth, are rendered using images with rich textures, and the animations are exported at a higher resolution. The objects are proportionally depicted with a realistic ratio. These animations incorporate more visual cues, including arrows, lines, and color. The effects for dynamic changes in the animation are more realistic. Each key knowledge point is enhanced with more realistic lighting rendering, and displayed with additional detail richness. Figure 2 illustrates a comparison between high-level and low-level conditions in visual detail.

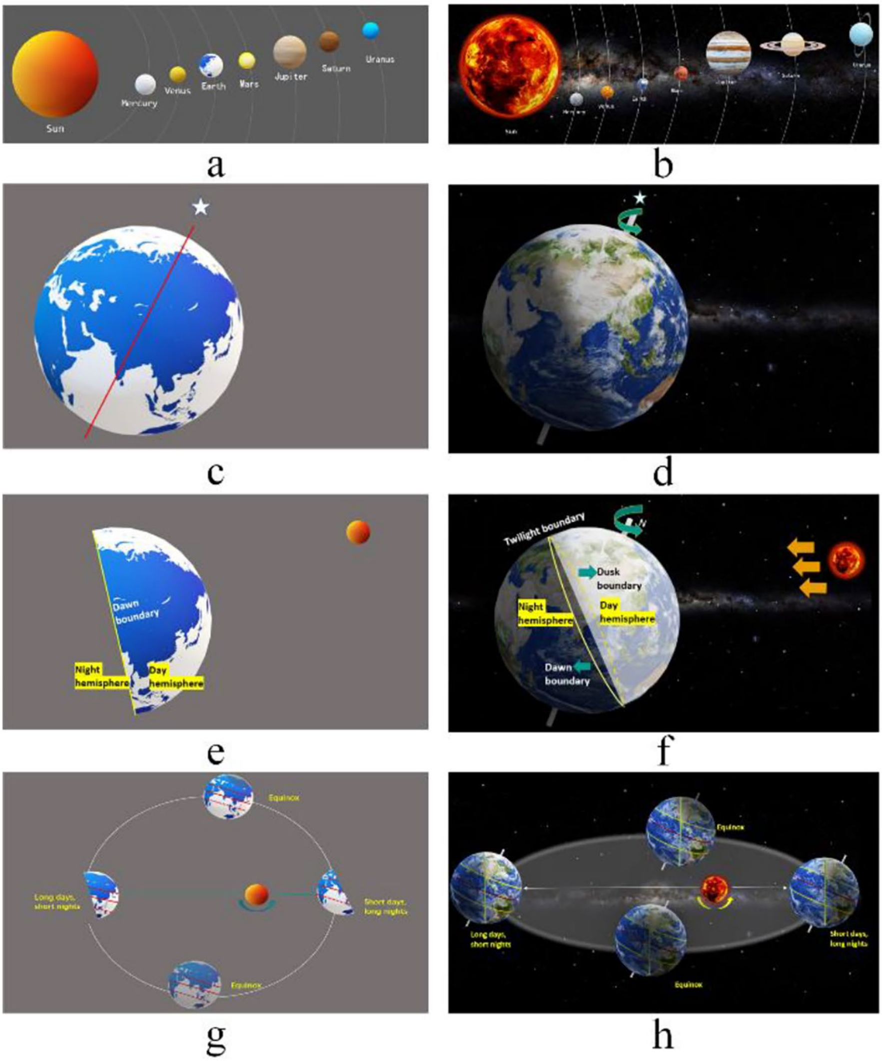


**Fig. 2** Comparison of detail richness examples between the high condition and the low condition (translated). Subfigures labeled as **a**, **c**, **e**, and **g** depict the low-level condition, while those labeled as **b**, **d**, **f**, and **h** represent the high-level condition

#### 2.2.2 Text load

Text load refers to the amount of text or written content displayed on a screen. In high text load conditions, the text includes keywords, names of key objects and phenomena, as well as text explanations related to the knowledge points. Conversely, in low text load conditions, the text only presents the keywords, object names, and phenomenon names. For instance, in the animation depicting the Earth's rotation with high text load, the text includes the following:

1. The name of the phenomenon: diurnal phenomenon, the alternation of day and night.
2. Explanation of the relationship between the sun and the Earth: the side facing the sun experiences daylight, while the opposite side experiences nighttime.
3. Clarification of the sun's role: the Earth itself does not emit light and is opaque; the sun serves as the sole light source.

In contrast, under low text load conditions, the textual descriptions only include the name of the phenomenon: diurnal phenomenon, the alternation of day and night.

Furthermore, the placement of text adheres to the contiguity principle, which suggests aligning on-screen text with corresponding graphics or animations (Moreno & Mayer, 1999a). For instance, the names of celestial bodies are positioned below them, while the explanation of the sun's role is located near the sun itself.

### 2.3 Dependent measures

The dependent variables studied were knowledge gains, cognitive load, learning experience, and emotions.

#### 2.3.1 Knowledge acquisition pre- and post-tests

Knowledge acquisition was assessed by administering a knowledge test, including factual and conceptual knowledge (Anderson et al., 2001). The present study employed this knowledge test as both pretest and posttest measures, which featured the identical content but a different randomized sequence of test items, to study pre-to-post changes on learning. The knowledge test, scoring criteria, point values, and answers for each question were taken from the exercise book and the teacher's edition textbook. The questions were designed to be consistent with the learning objectives set for the course, and their types and levels of difficulty were consistent with these objectives. This test consisted of five multiple-choice items and fifteen fill-in-the-blank questions (good internal consistency; Cronbach's alpha = 0.887), with a maximum score of 40 points. An example of a multiple-choice question was "The sun always rises in the east and sets in the west because of which of the following?" and the five response options were: (A) The Earth constantly rotates from west to east; (B) The Earth constantly rotates from east to west; (C) The Earth constantly rotates from north to south; and (D) The Earth constantly rotates from south to north. A correct answer was worth 2 points and an incorrect answer received 0 points. The maximum score for multiple-choice questions was therefore 10. An example of a fill-in-the-blank question was "The boundary between day and night on Earth is called ______". A correct answer was worth 2 points and an incorrect answer received 0 points. The maximum score for fill-in-the-blank questions was therefore 30. However, as the answers in some fill-in-the-blank questions consisted of two parts, a single correct answer was worth only one point, which could lead to an odd total score.

#### 2.3.2 Cognitive load

The cognitive load scale, comprising two questions, two questions, was adapted from (Paas & Van Merriënboer, 1994a; Paas & Merriënboer, 1994b). Participants were asked to mark their level of cognitive load using a Likert scale ranging from 1 to 9. The adapted scale demonstrated a high level of internal consistency with a reliability coefficient of 0.889.

#### 2.3.3 Learning experience

The learning experience scale was adapted for this study from learning experience questionnaires (Spain et al., 2008; Katuk et al., 2013; Hajhashemi et al., 2018). Developed as a Likert nine-point scale (ranging from 1, indicating strongly disagree, to 9, indicating strongly agree), this scale comprises three dimensions with a total of 5 items: Satisfaction (2 items), Trust (1 item), and Interest (2 items). The adapted scale demonstrated a high level of internal consistency, with a reliability coefficient of 0.889.

#### 2.3.4 Emotions

The Self-Assessment Manikin (SAM) (Bradley & Lang, 1994) is an instrument that uses graphic rating scales composed of images and characters to measure individual emotional responses induced by pictures (Backs et al., 2005), videos (Beege et al., 2018), and VR (Barreda-Ángeles et al., 2021; Somarathna et al., 2022). This scale has been widely used to measure emotional responses in various research contexts and has demonstrated high reliability and validity (Fekete et al., 2022). It consists of a set of three pictorial scales, each representing a different dimension of subjective experience: valence (unhappy to happy), arousal (relaxed to excited), and dominance (in-control to controlled). In this study, each scale has a set of nine pictorial figures (manikins), ranging from low to high on the corresponding dimension. Participants are asked to choose the figure that best represents their subjective experience at a given moment.

Participants in each experimental group were instructed to assess their emotional responses to screenshots depicting three essential concepts from the 3D animation they had been exposed to. The assessment of emotions included a set of nine questions (3 emotion dimensions × 3 animated learning resource screenshots) and demonstrated good reliability (Cronbach's alpha = 0.890).

### 2.4 Procedure

Figure 3 illustrates the experimental procedure. The course was conducted in a multimedia classroom where each student was equipped with a personal computer, and it lasted approximately 45 min. Before starting the course, students completed demographic information (e.g., gender) and the pretests. All four experimental groups were taught by the same instructor, following the same pedagogical method and procedure; the only difference was the 3D animation. The teaching duration was approximately 30 min. The 3D animated instructional resources were presented via a large display system. After the learning session, students completed knowledge acquisition posttests and subjective scales using their personal computers.

### 2.5 Data analysis

Based on the results of the Kruskal-Wallis test and visual assessments of the histograms, normal Q-Q plots, and box plots, it was determined that the observed values exhibited normal distribution across all four conditions. As a result, ANOVAs were deemed appropriate for further analysis. The dependent variables in this study included the pretest and posttest scores related to knowledge acquisition, as well as scores related to cognitive load, learning experience, and emotional responses.

## 3 Results

### 3.1 Knowledge acquisition

Table 1 presents an overview of the descriptive statistics across four groups. Table 2 shows an overview of main effect and interaction of independent variables.

The pretest results showed that participants had low prior knowledge of the materials to be learned, and there was no significant difference in prior knowledge across the four conditions ($F(3, 102)=0.014$, $p=.998$, $\eta^2 = 0.000$). This indicates that participants' familiarity with the knowledge were the same across all four conditions.

The results of the one-way ANOVA analysis indicated significant differences in the performance scores of students immediately after acquiring the knowledge across the four groups ($F(3,102)=3.322$, $p=.023$, $\eta^2 = 0.089$). Post hoc Tukey's HSD tests

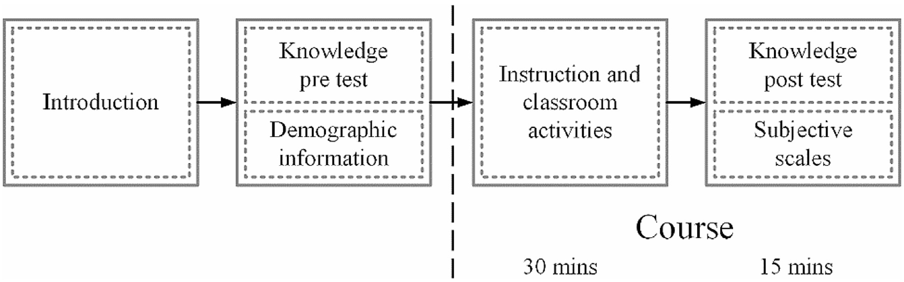


**Fig. 3** Visual overview of the study procedure

**Table 1** Means (M) and standard deviations (SD) of dependent variables across the four conditions

| Measures | | LDLT (*n*=27) | LDHT (*n*=26) | HDHT (*n*=26) | HDLT (*n*=27) |
|---|---|---|---|---|---|
| | | M (SD) | M (SD) | M (SD) | M (SD) |
| Knowledge acquisition pretest | | 7.30(5.14) | 7.38(4.58) | 7.46(5.39) | 7.56(3.84) |
| Knowledge acquisition posttest | | 17.56(7.12) | 14.77(7.28) | 18.15(6.42) | 20.89(7.48) |
| Cognitive load | Mental effort | 6.96(2.14) | 7.69(1.29) | 7.31(1.54) | 5.67(2.00) |
| | Mental load | 6.22(2.72) | 7.88(1.37) | 6.96(1.78) | 5.93(2.13) |
| Learning experience | Satisfaction | 6.65(1.81) | 6.46(2.08) | 7.62(1.58) | 7.74(1.53) |
| | Trust | 5.74(1.93) | 7.08(2.26) | 7.69(1.57) | 7.37(1.96) |
| | Interest | 6.78(2.12) | 6.71(2.18) | 7.62(1.60) | 7.67(1.70) |
| Emotions | Valence | 6.49(1.78) | 6.71(1.99) | 7.46(1.60) | 7.44(1.62) |
| | Arousal | 6.78(2.08) | 6.94(1.97) | 7.44(1.40) | 7.26(2.11) |
| | Dominance | 6.58(1.94) | 6.97(1.76) | 7.41(1.23) | 7.07(1.42) |

**Table 2** Main effect and interaction of independent variables

| Measures | | Detail richness (LDLT+LDHT) vs. (HDHT+HDLT) | | Text load (LDLT+HDLT) vs. (HDHT+LDHT) | | Detail richness * Text load | |
|---|---|---|---|---|---|---|---|
| | | *p* | $\eta^2$ | *p* | $\eta^2$ | *p* | $\eta^2$ |
| Knowledge acquisition posttest | | 0.016* | 0.055 | 0.048* | 0.038 | 0.098 | 0.000 |
| Cognitive load | Mental effort | 0.017* | 0.055 | 0.001** | 0.103 | 0.191 | 0.017 |
| | Mental load | 0.132 | 0.022 | 0.001** | 0.100 | 0.437 | 0.006 |
| Learning experience | Satisfaction | 0.001** | 0.095 | 0.650 | 0.002 | 0.929 | 0.000 |
| | Trust | 0.015* | 0.056 | 0.095 | 0.027 | 0.067 | 0.033 |
| | Interest | 0.018* | 0.054 | 0.875 | 0.000 | 0.984 | 0.000 |
| Emotions | Valence | 0.014* | 0.058 | 0.738 | 0.001 | 0.776 | 0.001 |
| | Arousal | 0.190 | 0.017 | 0.654 | 0.002 | 0.980 | 0.000 |
| | Dominance | 0.141 | 0.021 | 0.247 | 0.013 | 0.927 | 0.000 |

*$p$<.05;**$p$<.01

showed a significant difference between the HDLT and LDHT groups, with the HDLT group scoring higher (mean difference=6.120, $p$=.012). However, no significant differences were observed across other groups (See Fig. 4).

We conducted two-way ANOVAs for the four experimental groups. The main effects of detail richness (F(1,102)=5.948, $p$=.016, $\eta^2$ = 0.055) and text load (F(1,102)=4.018, $p$=.048, $\eta^2$ = 0.038) were both found to be significant. However, no interaction between the two independent variables was observed (F(1,102)=0.000, $p$=.985, $\eta^2$ = 0.000), as shown in Fig. 5.

## 3.2 Cognitive load

The results of the one-way ANOVA analysis indicated significant differences in cognitive load across the four groups (F(3,102)=6.507, $p$=.000, $\eta^2$ = 0.161). Post hoc Tukey's HSDs showed that the scores of HDLT differed significantly from the LDHT group, with the LDHT group having higher scores (mean difference=1.992, $p$=.000), and the scores of HDHT differed from the HDLT group, with the HDHT

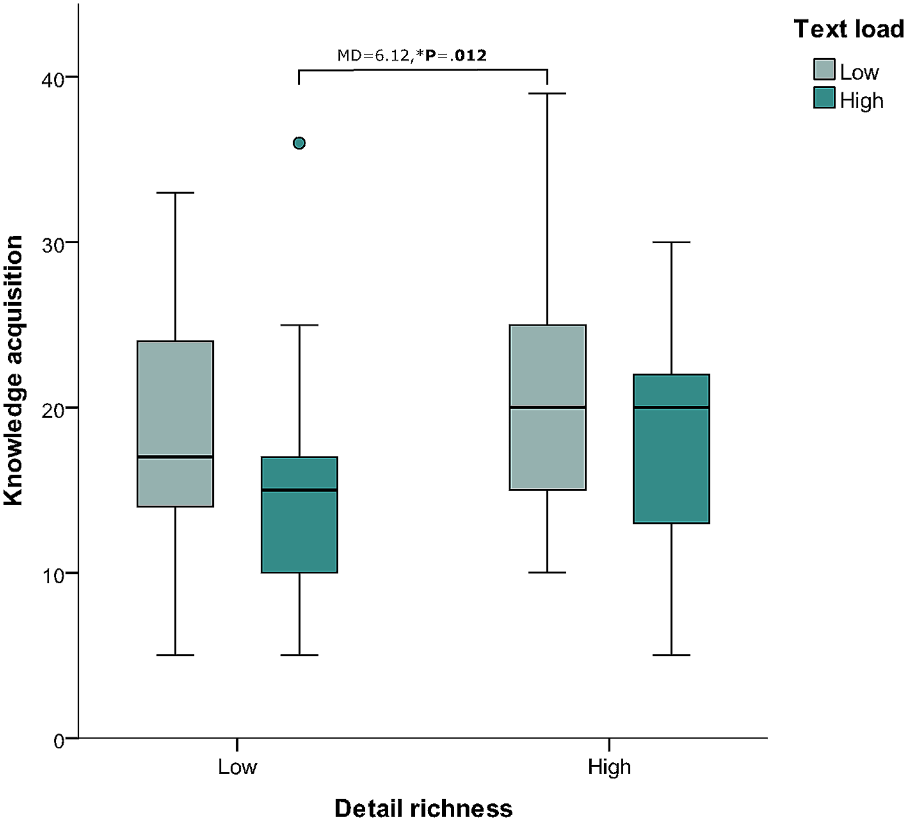


**Fig. 4** Differences in the knowledge acquisition across four conditions

group having higher scores (mean difference = 1.338, $p = .026$). However, there was no significant difference across these other groups (See Fig. 6).

The analysis of two-way ANOVAs showed that the main effects of the detail richness ($F(1,102) = 4.799$, $p = .031$, $\eta^2 = 0.045$) and text load ($F(1,102) = 14.655$, $p = .000$, $\eta^2 = 0.126$) on overall cognitive load were significant. However, there was no significant interaction between these two independent variables ($F(1, 102) = 0.046$, $p = .830$, $\eta^2 = 0.000$), as shown in Fig. 7.

A 2×2 MANOVA test was conducted for each scale to examine how visual fidelity and text load impacted mental effort and perceived task difficulty. The main effect of detail richness on mental effect was significant ($F(1,102) = 5.890$, $p = .017$, $\eta^2 = 0.055$), while it was not significant on perceived task difficulty. The main effect of text load was significant for both mental effect ($F(1,102) = 11.713$, $p = .001$, $\eta^2 = 0.103$) and perceived task difficulty ($F(1,102) = 11.275$, $p = .001$, $\eta^2 = 0.100$). However, no interaction was found between the two independent variables for each scale.

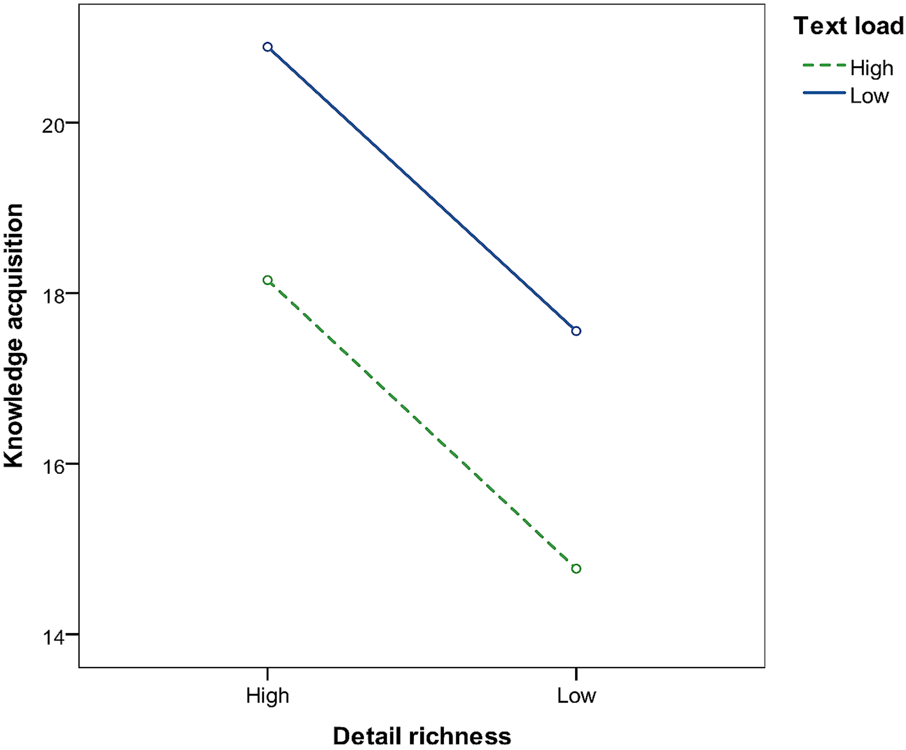


**Fig. 5** Interaction of the knowledge acquisition between detail richness and text load

### 3.3 Learning experience

The results of the one-way ANOVA analysis indicated significant differences in learning experience across four groups (F(3,102) = 3.606, $p$ = .016, $\eta^2$ = 0.096). Post hoc Tukey's HSDs showed that the scores of LDLT differed from the HDLT group, with the HDLT group having higher scores (mean difference = 1.204, $p$ = .034). However, there was no significant difference across other groups (See Fig. 8).

The analysis of two-way ANOVAs on overall learning experience showed that the main effect of detail richness (F(1,102) = 10.021, $p$ = .002, $\eta^2$ = 0.089) was significant, while the main effect of text load was not significant. Besides, we did not find an interaction between the two independent variables (F(1,102) = 0.504, $p$ = .480, $\eta^2$ = 0.005).

A 2 × 2 MANOVA test was conducted for each scale to examine how detail richness and text load impacted satisfaction, trust, and interest. The main effects of detail richness on satisfaction (F(1,102) = 10.764, $p$ = .001, $\eta^2$ = 0.095), trust (F(1,102) = 6.075, $p$ = .015, $\eta^2$ = 0.056), and interest (F(1,102) = 5.775, $p$ = .018, $\eta^2$ = 0.054) were significant. However, the main effects of text load on these scales were not significant. Besides, we did not find an interaction between the two independent variables for each scale.

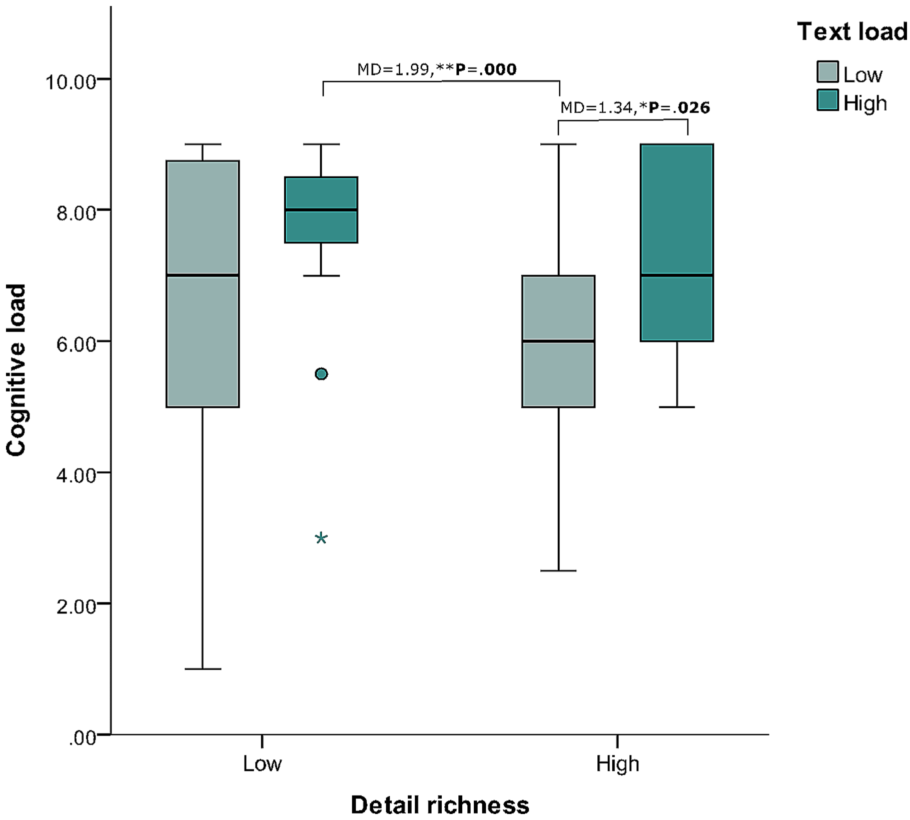


**Fig. 6** Differences in the overall cognitive load across four conditions

### 3.4 Emotions

The results of the one-way ANOVA analysis indicated that there were no significant differences in learning experience across four groups ($F(3,102)=1.407$, $p=.245$, $\eta^2 = 0.040$).

The analysis of two-way ANOVAs showed that the main effects of detail richness and text load on overall emotions were not significant. Besides, we did not find an interaction between the two independent variables.

A 2×2 MANOVA test was conducted for each scale to examine how detail richness and text load impacted valence, arousal, and dominance. The main effect of detail richness on valence was significant ($F(1,102)=6.265$, $p=.014$, $\eta^2 = 0.058$), while the other two scales were not significant. The main effect of text load on all three scales of emotions was not significant. Additionally, we did not find an interaction between the two independent variables for each scale.

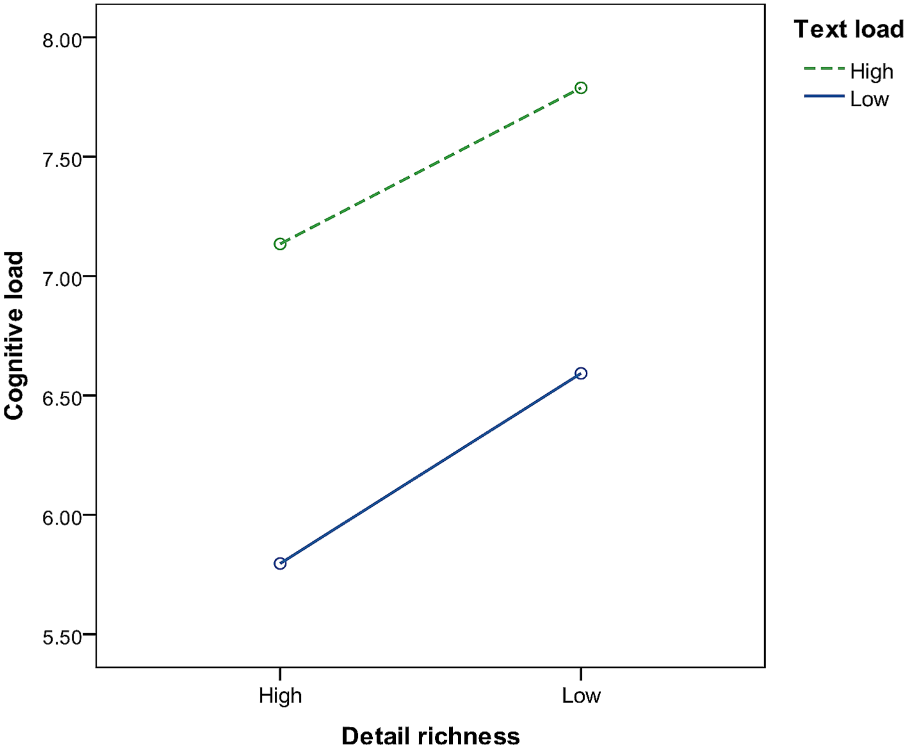


**Fig. 7** Interaction of the overall cognitive load between detail richness and text load

## 4 General discussion

### 4.1 RQ1: How do detail richness (low/high) and text load (low/high) affect the acquisition of knowledge?

First, the results indicated that the presence of rich detail in 3D animation has a positive impact on learning performance. The groups exposed to high detailed imagery achieved significantly higher average scores than those exposed to less detailed imagery. One possible explanation for this is that enhanced visibility assists learners in creating mental imagery, leading to an improvement in their learning performance. Previous research has established that the use of appropriate generative learning strategies within multimedia learning environments effectively engages learners in cognitive processes, thereby enhancing learning outcomes (Mayer, 2014a; Fiorella & Mayer, 2016; Leopold et al., 2019; Schmidgall et al., 2019; de Koning et al., 2020). Among these strategies, imagery techniques, often referred to as mental practice, find widespread use during the learning process (Feltz & Landers, 1983). The learning tasks in this study required the construction and use of mental imagery. For example, when students seek to comprehend the concept of "what does the Earth look like when facing the sun versus when facing away from the sun," they must mentally visualize the varying positions and appearances of the Earth. While experts can gen-

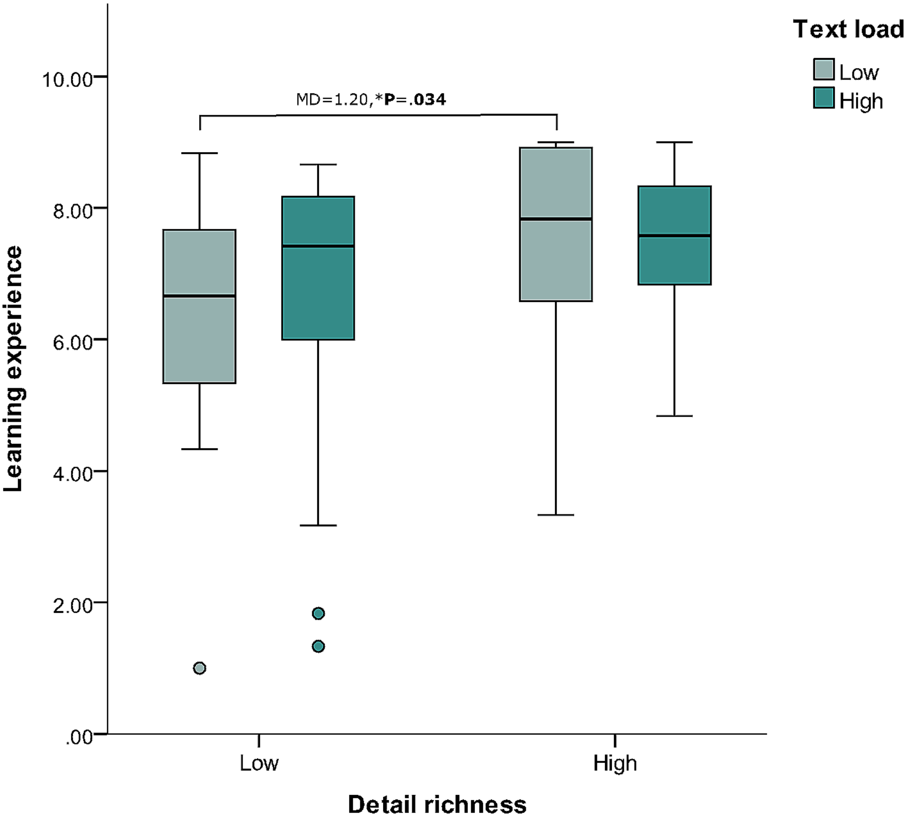


**Fig. 8** Differences in the learning experience across four conditions

erate their own mental images, the less experienced students rely on the assistance of visualization to facilitate the construction of mental imagery (Mayer & Gallini, 1990). Rich details may help such less experienced students facilitate the externalization of mental imagery. It is worth noting that visualization plays an essential role in facilitating the construction of mental images. Sibley et al. have observed that the visibility of learning materials significantly influences the use of imagery strategies, with visibility leading to a positive impact on learning through imagery (Sibley et al., 2022). Building upon the analysis of findings on detail richness, we believe that the inclusion of rich visual details in instructional materials serves as scaffolding, making it easier for students to engage in mental imagery and, subsequently, positively influencing their learning experiences. In essence, enhanced visibility leads to a more beneficial influence on learning that requires imagery.

Furthermore, animations may contain complicated scenarios or sequences, making it difficult for learners to identify important information and establish meaningful connections between pieces of information. The use of rich visual cues, such as color highlights and arrows, can help learners pay attention to critical components and facilitate the integration of elements within the animation (de Koning et al., 2011; Lin & Atkinson, 2011; Boucheix et al., 2013).

Second, concerning text load, our findings have shown that adding a small amount of textual description to 3D animations can significantly enhance learning outcomes. The multimedia principle suggests active construction, which emphasizes that learners themselves can mentally construct graphical and verbal representations of the material and seamlessly integrate them together (Butcher, 2014). When students engage in mental visualization and make connections between the graphical and verbal materials, their cognitive processing of the information is enhanced. This process promotes deeper understanding and meaningful engagement with the material, resulting in improved learning outcomes.

### 4.2 RQ2: How do detail richness and text load affect the cognitive load?

First, the results indicated that the high detail richness groups experienced significantly lower levels of cognitive load than the low detail richness groups. When learning involves mental imagery of a process, we have observed that reducing detail richness can actually increase cognitive load. Some researchers argue that abstraction is beneficial for learning and discourage realistic visualization that includes irrelevant details, providing only necessary visual information (Renkl & Scheiter, 2017). This perspective posits that abstract visualization helps students focus their attention on relevant information and experience lower cognitive load. However, our research revealed that visualization with fewer details does not always lead to reduced cognitive load; it may also increase students' perceived task difficulty and mental effort. This may be because geography learning in our study requires mental imagery, and visualization with fewer details may not effectively support beginners in forming accurate mental images, which then hinders subsequent knowledge processing. Taken together, from the perspective of cognitive load theory, when visual materials lack details, learners may need to invest more cognitive resources to mentally construct the missing details, leading to higher cognitive load and lower learning performance.

Second, we observed that the high text load groups experienced significantly higher levels of cognitive load than the low text load groups. Viewed through the lens of cognitive load theory (Sweller, 2011), the animation alone may lead to increased cognitive load. When textual descriptions are integrated, they serve as valuable aids in directing learner attention and optimizing cognitive load. On-screen text provides clear guidance, facilitating learners' mental processes and reducing the cognitive demands required for comprehension. This guidance allows learners to devote more of their attention to meaningful learning activities. Furthermore, while the visual aspects of 3D animations convey important information, the inclusion of concise textual descriptions complements and reinforces the visual content. The integration of text and detailed imagery ensures that learners receive the necessary information from multiple sources, enhancing their knowledge acquisition, learning experience and emotions. However, it is important to consider that learners already receive audio information through the auditory channel and visual information through the visual channel. Therefore, including textual information in the visual channel may further increase the cognitive load according to the redundancy principle (Kalyuga et al., 2004; Gerjets et al., 2009; Mayer, 2014b, 2020). Our research suggests that the textual descriptions included in 3D animations should be concise, containing only keywords and names of key objects. Increasing the amount of textual description, on the other hand, will only increase the cognitive load, resulting in reduced learning performance and a less satisfying learning experience. Therefore, we recommend including a small amount of text in the 3D animation and place it according to the contiguity principle (Moreno & Mayer, 1999a).

### 4.3 RQ3: How do detail richness and text load affect the learning experience and emotion?

The results indicated that the presence of rich detail in 3D animation has a positive impact on learning experience and emotion. The high-detail groups had significantly higher learning experience scores than the low-detail groups. Although there were no significant differences in emotions between the high and low visual detail groups, the high visual detail groups reported higher levels of pleasure, arousal, and dominance. The significant main effects of detail richness on satisfaction, trust, interest, and valence suggest that detailed imagery evokes a positive emotional experience for students during learning. Previous research suggests that emotions play a significant role in impacting how learners process information and their motivation to learn, consequently affecting overall learning outcomes (Pekrun & Linnenbrink-Garcia, 2014). When the learning objective requires the use of mental imagery, incorporating rich detail can facilitate the externalization of mental imagery, thereby enhancing the acquisition of knowledge related to process understanding, as well as improving the emotional and experiential aspects of learning. Furthermore, we did not find any significant differences between the high and low text load groups regarding the aspects of learning experience and emotions. Taken together, these findings shed light on the critical role of rich detail in 3D animation in shaping the learning experience and emotions of students. Additionally, the study provides insights into the relatively minor role of text load compared to visual detail in influencing students' learning experiences and emotions. This not only contributes to our understanding of how detail richness is connected to learning objectives related to process understanding but also addresses the gap in the existing literature regarding the impact of detailed visualizations on the learning experience and emotions.

### 4.4 RQ4: Do detail richness and text load interact in their impact on learning?

No significant interaction effect between detail richness and text load was found for learning performance, cognitive load, learning experience, and emotions. High detailed imagery groups consistently outperformed low detailed imagery groups in terms of performance and learning experience. Similarly, low text load groups consistently outperformed high text load groups in terms of performance and learning experience. Notably, the group exposed to rich detailed imagery and low text load showed the highest level of learning performance. In addition, students in the high detailed imagery, low text load group reported improved learning experiences and more positive emotions related to learning. This inconsistency between detail richness and text load of 3D animation may be important for learning the process, particularly in the context of geography learning.

In terms of knowledge acquisition, both the main effects of detail richness and text load were significant, with detail richness showing a greater significance than text load. This suggests that detail richness had a more significant impact on knowledge acquisition than text load, indicating its critical role in shaping performance. Regarding learning experience and emotions, detail richness had a significant effect, while the main effect of text load was not significant. This indicates that detail richness independently influenced learning experience and emotions, regardless of text load. Conversely, text load had no significant effect on learning experience and emotions. In terms of cognitive load, both the main effects of detail richness and text load were significant, with text load showing a greater significance than detail richness. This indicates that text load had a stronger impact on cognitive load than detail richness. Therefore, to minimize unnecessary cognitive load, we recommend increasing the detail of 3D animations and using concise text when demonstrating knowledge about the processes or systems through 3D animations.

### 4.5 Implications for the design of dynamic visualization

First, this study can serve as an illustrative example of effective visualization design for learning about processes. The findings indicate that the use of realistic and detailed visualizations can significantly enhance the performance. In this study, the focus was on a geographic process that encompassed not only the factual knowledge about entities like Earth and Sun but also the conceptual knowledge about their interrelationships. The use of realistic and detailed visualizations provided students with a solid mental scaffold, allowing for better knowledge acquisition and an improved learning experience. Conversely, simplified visualizations in this context did not provide adequate mental scaffolding, leading to inadequate mental imagery processing and high cognitive load. It is crucial to note that maintaining a consistent level of realism and detail across the 3D animations is essential. In addition to the realistic rendering of objects, factors such as scale, environment, and animated visual cues should also be considered.

Another important implication for the design of dynamic visualizations is the effective use of textual descriptions. The research findings highlight that the inclusion of key object names and relevant keywords can have a significant positive impact on learning performance and experience. By including these essential textual elements, students' attention can be appropriately directed without overwhelming their cognitive workload. It is crucial to ensure that the verbal messages displayed are concise, using only a few words to convey the necessary information. When developing 3D animation learning resources, complex and redundant textual descriptions should be avoided, as they have the potential to increase students' extraneous cognitive load, leading to cognitive overload and diminished learning outcomes. Furthermore, it is crucial to adhere to the contiguity principle when designing dynamic visualizations. This principle emphasizes the importance of avoiding spatial separation between words and visual elements within the learning materials. By closely aligning the text with the depicted objects, a seamless integration of text and visuals can be achieved, facilitating better comprehension and understanding. Therefore, when designing dynamic visualizations, careful consideration should be given to the concise and strategically placed textual descriptions to ensure optimal learning outcomes.

## 5 Conclusion, limitations and future directions

This study contributes insights into the impact of detailed visualizations and text load on the performance, experience, and emotions of high school students in geography learning. The results showed that students performed better in acquiring knowledge when exposed to a condition characterized by high detailed imagery and low text load. Conversely, both low detail richness and high text load independently resulted in increased cognitive load. Furthermore, a significant impact of detail richness on the dimensions of pleasure and arousal was observed, with higher levels of detail correlating with greater values in these dimensions. Taken together, the findings emphasize the role of detailed visualization in facilitating the externalization of mental imagery and cognitive processing when learning goals require its involvement. The results also indicated the effectiveness of high-level detailed imagery and concise text, suggesting alignment of realism with task demands and pedagogical goals. The findings hold significant implications for educators and practitioners aiming to improve instructional design for high school learners.

However, it is important to note that the present study has a few limitations that should be addressed in future investigations. First, our focus was on measuring conceptual learning through pre- and post-test changes on a multiple-choice and fill-in-the-blank question test administered before and after the intervention. Yet, we did not examine the transfer of learning, which involves the application of acquired knowledge, skills, or strategies in new contexts. Transfer is important in learning as it enables learners to apply their knowledge effectively in diverse settings, thereby increasing their overall competence and confidence. Another limitation of the current study is its focus on the effect patterns of learning about the process of 3D animations, which lacks user interaction. When it comes to VR, individuals can interact and navigate within immersive environments, which adds additional dimensions to consider. Recognizing VR as a promising environment for facilitating learning, our future research will be dedicated to specifically investigating the impact of detailed visualization within VR environments. Third, as learners play the central role in the learning process, their individual differences inevitably impact the effectiveness of detailed visualization. According to the findings in (Zhou et al., 2022), individuals have different levels of mental imagery, which consequently affects their performance. Therefore, it is crucial to thoroughly investigate the interplay of the different levels of learners' mental imagery and detailed visualization to gain a deeper understanding of their impact on the learning. Finally, it is important to note that the duration of the study conducted in this research was relatively short-term. As a result, further investigation is needed to explore whether the positive outcomes and findings observed in this study can be sustained and replicated over longer periods of learning. This will provide valuable insights into the long-term implications and practical applications of the findings, allowing for more informed decision-making in educational settings.

When designing dynamic visualizations to facilitate learning about processes, it is recommended to use detailed visual representations and provide concise textual descriptions for the best learning outcomes and experience. First, when the instructional objectives of a course or application involve process-based learning, the recommendation is to design realistic and detailed visualizations. Consistency in incorporating detailed visuals throughout 3D animations is also crucial. Second, it is important to incorporate keywords into textual descriptions. Additionally, closely aligns text with visual elements in learning materials can promote the integration of information in the brain. By shedding light on the connection between detail richness and process understanding, this research enhances our understanding of designing visual elements and text for 3D animations for geography learning.

**Authors contributions** Yun Zhou and Tao Xu contributed to the study conception and design. Material preparation and data collection and analysis were performed by Fanqi Yi, Bingyu Dong, Guangli Zhang, and Yi Zhang. The first draft of the manuscript was written by Yun Zhou. All authors read and approved the final manuscript.

**Funding** This work was supported by National Key R&D Program of China under Grant [2022YFC3303600]; the Research Projects of the National Natural Science Foundation of China under Grant [62077036] and [62377039].

**Data availability** Data will be available from the corresponding author upon reasonable request.

**Code Availability** The application used in this study will be available from the corresponding author upon reasonable request.

## Declarations

**Ethics approval** This research was approved by the Ethics Committee of the Faculty of Education at Shaanxi Normal University.

**Consent to participate** The school and students' parents were informed about this investigation and asked for an informed consent agreement.

**Conflict of interest** The authors declare no competing interests relevant to the content of this article.

## References


Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J. D., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of educational objectives* (Complete ed.). Longman.

Ardaç, D., & Unal, S. (2008). Does the amount of on-screen text influence student learning from a multimedia-based instructional unit? *Instructional Science*, *36*(1), 75–88. https://doi.org/10.1007/s11251-007-9035-4.

Awidi, I. T., & Paynter, M. (2019). The impact of a flipped classroom approach on student learning experience. *Computers & Education*, *128*, 269–283. https://doi.org/10.1016/j.compedu.2018.09.013.

Backs, R. W., Silva, S. P., da, & Han, K. (2005). A comparison of younger and older adults' Self-Assessment Manikin ratings of affective pictures. *Experimental Aging Research*, *31*(4), 421–440. https://doi.org/10.1080/03610730500206808.

Barreda-Ángeles, M., Serra-Blasco, M., Trepat, E., Pereda-Baños, A., Pàmias, M., Palao, D., Goldberg, X., & Cardoner, N. (2021). Development and experimental validation of a dataset of 360°-videos for facilitating school-based bullying prevention programs. *Computers & Education*, *161*, 104065. https://doi.org/10.1016/j.compedu.2020.104065.

Beege, M., Schneider, S., Nebel, S., Häßler, A., & Rey, G. D. (2018). Mood-affect congruency. Exploring the relation between learners' mood and the affective charge of educational videos. *Computers & Education*, *123*, 85–96. https://doi.org/10.1016/j.compedu.2018.05.001.

Belenky, D. M., & Schalk, L. (2014). The effects of Idealized and grounded materials on learning, transfer, and interest: An Organizing Framework for Categorizing External Knowledge representations. *Educational Psychology Review*, *26*(1), 27–50. https://doi.org/10.1007/s10648-014-9251-9.

Berney, S., & Bétrancourt, M. (2016). Does animation enhance learning? A meta-analysis. *Computers & Education*, *101*, 150–167. https://doi.org/10.1016/j.compedu.2016.06.005.

Birbara, N. S., & Pather, N. (2021). Real or not real: The impact of the Physical Fidelity of virtual Learning resources on learning anatomy. *Anatomical Sciences Education*, *14*(6), 774–787. https://doi.org/10.1002/ase.2022.

Boucheix, J. M., Lowe, R. K., Putri, D. K., & Groff, J. (2013). Cueing animations: Dynamic signaling aids information extraction and comprehension. *Learning and Instruction*, *25*, 71–84. https://doi.org/10.1016/j.learninstruc.2012.11.005.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9.

Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology*, *98*, 182–197. https://doi.org/10.1037/0022-0663.98.1.182.

Butcher, K. R. (2014). The multimedia principle. In *The Cambridge handbook of multimedia learning, 2nd ed* (pp. 174–205). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.010.

Castro-Alonso, J. C., Ayres, P., & Sweller, J. (2019). Instructional Visualizations, Cognitive Load Theory, and Visuospatial Processing. In J. C. Castro-Alonso (Ed.), *Visuospatial Processing for Education in Health and Natural Sciences* (pp. 111–143). Springer International Publishing. https://doi.org/10.1007/978-3-030-20969-8_5.

Chen, C. H. (2020). Impacts of augmented reality and a digital game on students' science learning with reflection prompts in multimedia learning. *Educational Technology Research and Development*, *68*(6), 3057–3076. https://doi.org/10.1007/s11423-020-09834-w.

de Koning, B., Tabbers, H., Rikers, R., & Paas, F. (2011). Improved effectiveness of Cueing by Self-explanations when learning from a Complex Animation. *Applied Cognitive Psychology*, *25*, 183–194. https://doi.org/10.1002/acp.1661.

de Koning, B., Rop, G., & Paas, F. (2020). Effects of spatial distance on the effectiveness of mental and physical integration strategies in learning from split-attention examples. *Computers in Human Behavior*, *110*, https://doi.org/10.1016/j.chb.2020.106379.

Di Natale, A. F., Repetto, C., Riva, G., & Villani, D. (2020). Immersive virtual reality in K-12 and higher education: A 10-year systematic review of empirical research. *British Journal of Educational Technology*, *51*(6), 2006–2033. https://doi.org/10.1111/bjet.13030.

El-Nahass, N. M., & Abdellatif, M. A. I. (2021). Students' learning difficulties in Geography in (K-12) classrooms and instructional interventions according to Representational systems of personality. *American Journal of Education and Information Technology*, *5*(1). https://doi.org/10.11648/j.ajeit.20210501.17.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146.

Fekete, A., Pelowski, M., Specker, E., Brieber, D., Rosenberg, R., & Leder, H. (2022). *The Vienna Art Picture System (VAPS): A data set of 999 paintings and subjective ratings for art and aesthetics research*. No Pagination Specified-No Pagination Specified. https://doi.org/10.1037/aca0000460.

Feltz, D. L., & Landers, D. M. (1983). The effects of mental practice on motor skill learning and performance: A meta-analysis. *Journal of Sport Psychology*, *5*(1), 25–57.

Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote Generative Learning. *Educational Psychology Review*, *28*(4), 717–741. https://doi.org/10.1007/s10648-015-9348-9.

Gerjets, P., Scheiter, K., Opfermann, M., Hesse, F. W., & Eysink, T. H. S. (2009). Learning with hypermedia: The influence of representational formats and different levels of learner control on performance and learning behavior. *Computers in Human Behavior*, *25*(2), 360–370. https://doi.org/10.1016/j.chb.2008.12.015.

Guerra-Tamez, C. R. (2023). The impact of immersion through virtual reality in the Learning experiences of Art and Design students: The mediating effect of the Flow experience. *Education Sciences*, *13*(2). https://doi.org/10.3390/educsci13020185.

Hajhashemi, K., Caltabiano, N., Anderson, N., & Tabibzadeh, S. (2018). Multiple intelligences, motivations and learning experience regarding video-assisted subjects in a Rural University. *International Journal of Instruction*, *11*, 167–182. https://doi.org/10.12973/iji.2018.11112a.

Han, Z. M., Huang, C. Q., Yu, J. H., & Tsai, C. C. (2021). Identifying patterns of epistemic emotions with respect to interactions in massive online open courses using deep learning and social network analysis. *Computers in Human Behavior*, *122*, 106843. https://doi.org/10.1016/j.chb.2021.106843.

Harrington, M. C. R. (2012). The virtual Trillium Trail and the empirical effects of Freedom and Fidelity on discovery-based learning. *Virtual Reality*, *16*(2), 105–120. https://doi.org/10.1007/s10055-011-0189-7.

Hegarty, M. (2004). Dynamic visualizations and learning: Getting to the difficult questions. *Learning and Instruction*, *14*(3), 343–351. https://doi.org/10.1016/j.learninstruc.2004.06.007.

Hegarty, M., Smallman, H., & Stull, A. (2012). Choosing and using Geospatial displays: Effects of Design on Performance and Metacognition. *Journal of Experimental Psychology Applied*, *18*, 1–17. https://doi.org/10.1037/a0026625.

Joo, H., Park, J., & Kim, D. (2021). Visual representation fidelity and self-explanation prompts in multi-representational adaptive learning. *Journal of Computer Assisted Learning*, *37*(4), 1091–1106. https://doi.org/10.1111/jcal.12548.

Kablan, Z., & Erden, M. (2008). Instructional Efficiency of Integrated and separated text with animated presentations in computer-based science instruction. *Computers & Education*, *51*(2), 660–668. https://doi.org/10.1016/j.compedu.2007.07.002.

Kalyuga, S., Chandler, P., & Sweller, J. (2004). When redundant On-Screen text in Multimedia technical instruction can interfere with learning. *Human Factors*, *46*(3), 567–581. https://doi.org/10.1518/hfes.46.3.567.50405.

Kang, H. (2021). Sample size determination and power analysis using the G*Power software. *Journal of Educational Evaluation for Health Professions*, *18*, 17. https://doi.org/10.3352/jeehp.2021.18.17.

Katuk, N., Kim, J., & Ryu, H. (2013). Experience beyond knowledge: Pragmatic e-learning systems design with learning experience. *Computers in Human Behavior*, *29*(3), 747–758. https://doi.org/10.1016/j.chb.2012.12.014.

Lee, A. L., Wong, E., K. W., & Fung, C. C. (2010). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Computers & Education*, *55*(4), 1424–1442. https://doi.org/10.1016/j.compedu.2010.06.006.

Leopold, C., Mayer, R. E., & Dutke, S. (2019). The power of imagination and perspective in learning from Science text. *Journal of Educational Psychology*, *111*(5), 793–808. https://doi.org/10.1037/edu0000310.

Li, L. Y., Huang, W. L., & Tsai, C. C. (2019). Development and evaluation of a video playing interface with headings and table of contents. *Interactive Learning Environments*, *28*, 1–16. https://doi.org/10.1080/10494820.2019.1707693.

Lin, L., & Atkinson, R. K. (2011). Using animations and visual cueing to support learning of scientific concepts and processes. *Computers & Education*, *56*(3), 650–658. https://doi.org/10.1016/j.compedu.2010.10.007.

Lowe, R. K., & Schnotz, W. (2014). Animation principles in Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd ed., pp. 513–546). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.026.

Mayer, R. E. (2003). The Promise of Multimedia Learning: Using the same Instructional Design methods across different media. *Learning and Instruction*, *13*(2), 125–139.

Mayer, R. E. (2014a). Cognitive theory of multimedia learning. In *The Cambridge handbook of multimedia learning, 2nd ed* (pp. 43–71). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.005.

Mayer, R. E. (Ed.). (2014b). *The Cambridge Handbook of Multimedia Learning* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.

Mayer, R. E. (2020, July 9). *Multimedia Learning*. Higher Education from Cambridge University Press; Cambridge University Press. https://doi.org/10.1017/9781316941355.

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology*, *83*, 484–490. https://doi.org/10.1037/0022-0663.83.4.484.

Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, *84*, 444–452. https://doi.org/10.1037/0022-0663.84.4.444.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, *82*(4), 715–726. https://doi.org/10.1037/0022-0663.82.4.715.

Mayer, R. E., & Johnson, C. I. (2008). Revising the redundancy principle in multimedia learning. *Journal of Educational Psychology*, *100*, 380–386. https://doi.org/10.1037/0022-0663.100.2.380.

Menendez, D., Rosengren, K. S., & Alibali, M. W. (2020). Do details bug you? Effects of perceptual richness in learning about biological change. *Applied Cognitive Psychology*, *34*(5), 1101–1117. https://doi.org/10.1002/acp.3698.

Moreno, R., & Mayer, R. E. (1999a). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, *91*, 358–368. https://doi.org/10.1037/0022-0663.91.2.358.

Moreno, R., & Mayer, R. E. (1999b). Multimedia-supported metaphors for meaning making in Mathematics. *Cognition and Instruction*, *17*(3), 215–248.

Moreno, R., & Mayer, R. E. (2002). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology*, *94*, 156–163. https://doi.org/10.1037/0022-0663.94.1.156.

Paas, F. G. W. C., & Van Merriënboer, J. J. G. (1994a). 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.

Paas, F. G. W. C., & Van Merriënboer, J. J. G. (1994b). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology*, *86*, 122–133. https://doi.org/10.1037/0022-0663.86.1.122.

Pekrun, R., & Linnenbrink-Garcia, L. (Eds.). (2014). International Handbook of Emotions in Education (1st ed.). Routledge. https://doi.org/10.4324/9780203148211.

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and Student Engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 259–282). Springer US. https://doi.org/10.1007/978-1-4614-2018-7_12.

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second-language multimedia learning environment. *Journal of Educational Psychology*, *90*, 25–36. https://doi.org/10.1037/0022-0663.90.1.25.

Ploetzner, R., Berney, S., & Bétrancourt, M. (2020). A review of learning demands in instructional animations: The educational effectiveness of animations unfolds if the features of change need to be learned. *Journal of Computer Assisted Learning*, *36*(6), 838–860. https://doi.org/10.1111/jcal.12476.

Rau, M. A. (2017). Conditions for the effectiveness of multiple visual representations in enhancing STEM learning. *Educational Psychology Review*, *29*(4), 717–761. https://doi.org/10.1007/s10648-016-9365-3.

Rau, M. A. (2020). Comparing multiple theories about learning with physical and virtual representations: Conflicting or complementary effects? *Educational Psychology Review*, *32*(2), 297–325. https://doi.org/10.1007/s10648-020-09517-1.

Renkl, A., & Scheiter, K. (2017). Studying visual displays: How to Instructionally Support Learning. *Educational Psychology Review*, *29*(3), 599–621. https://doi.org/10.1007/s10648-015-9340-4.
Rickey, M. G., & Bein, F. L. (1996). Students' learning difficulties in Geography and teachers' interventions: Teaching cases from K-12 classrooms. *Journal of Geography*, *95*(3), 118–125. https://doi.org/10.1080/00221349608978703.
Scheiter, K., Gerjets, P., Huk, T., Imhof, B., & Kammerer, Y. (2009). The effects of realism in learning with dynamic visualizations. *Learning and Instruction*, *19*(6), 481–494. https://doi.org/10.1016/j.learninstruc.2008.08.001.
Schmidgall, S. P., Eitel, A., & Scheiter, K. (2019). Why do learners who draw perform well? Investigating the role of visualization, generation and externalization in learner-generated drawing. *Learning and Instruction*, *60*, 138–153. https://doi.org/10.1016/j.learninstruc.2018.01.006.
Sibley, L., Fiorella, L., & Lachner, A. (2022). It's better when I see it: Students benefit more from open-book than closed-book teaching. *Applied Cognitive Psychology*, *36*(6), 1347–1355. https://doi.org/10.1002/acp.4017.
Skulmowski, A. (2022a). Is there an optimum of realism in computer-generated instructional visualizations? *Education and Information Technologies*, *27*(7), 10309–10326. https://doi.org/10.1007/s10639-022-11043-2.
Skulmowski, A. (2022b). Realistic visualizations can aid transfer performance: Do distinctive shapes and descriptive labels contribute towards learning? *Journal of Computer Assisted Learning*, *38*(3), 681–691. https://doi.org/10.1111/jcal.12640.
Skulmowski, A. (2022c). When color coding backfires: A guidance reversal effect when learning with realistic visualizations. *Education and Information Technologies*, *27*(4), 4621–4636. https://doi.org/10.1007/s10639-021-10796-6.
Skulmowski, A. (2023). Realistic details impact learners independently of split-attention effects. *Cognitive Processing*. https://doi.org/10.1007/s10339-022-01123-z.
Skulmowski, A., & Rey, G. D. (2021a). Realism as a retrieval cue: Evidence for concreteness-specific effects of realistic, schematic, and verbal components of visualizations on learning and testing. *Human Behavior and Emerging Technologies*, *3*(2), 283–295. https://doi.org/10.1002/hbe2.209.
Skulmowski, A., & Rey, G. D. (2021b). Visualizing pathogens: Disfluent shapes of pathogens increase their perceived complexity and danger while realism and disfluency boost the credibility of visualizations. *Human Behavior and Emerging Technologies*, *3*(2), 316–323. https://doi.org/10.1002/hbe2.232.
Smallman, H. S., & John, M. S. (2005). *Naive realism*: Misplaced faith in realistic displays. *Ergonomics in Design: The Quarterly of Human Factors Applications*, *13*(3), 6–13. https://doi.org/10.1177/106480460501300303.
Somarathna, R., Bednarz, T., & Mohammadi, G. (2022). Virtual reality for emotion Elicitation– A review. *IEEE Transactions on Affective Computing*, 1–21. https://doi.org/10.1109/TAFFC.2022.3181053.
SpainR. D., BustamanteE. A., & BlissJ. P. (2008). Towards an empirically developed scale for System Trust: Take two. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *52*(19), 1335–1339. https://doi.org/10.1177/154193120805201907.
Sweller, J. (2011). CHAPTER TWO - Cognitive Load Theory. In J. P. Mestre & B. H. Ross (Eds.), *Psychology of Learning and Motivation* (Vol. 55, pp. 37–76). Academic Press. https://doi.org/10.1016/B978-0-12-387691-1.00002-8.
Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 years later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5.
Tarchi, C., Zaccoletti, S., & Mason, L. (2021). Learning from text, video, or subtitles: A comparative analysis. *Computers & Education*, *160*, 104034. https://doi.org/10.1016/j.compedu.2020.104034.
Waller, D., Knapp, D., & Hunt, E. (2001). Spatial representations of virtual mazes: The role of visual fidelity and individual differences. *Human Factors*, *43*(1), 147–158. https://doi.org/10.1518/001872001775992561.
Wu, C. F., & Chiang, M. C. (2013). Effectiveness of applying 2D static depictions and 3D animations to orthographic views learning in graphical course. *Computers & Education*, *63*, 28–42. https://doi.org/10.1016/j.compedu.2012.11.012.
Yue, C. L., Bjork, E. L., & Bjork, R. A. (2013). Reducing verbal redundancy in multimedia learning: An undesired desirable difficulty? *Journal of Educational Psychology*, *105*, 266–277. https://doi.org/10.1037/a0031971.

Zhou, Y., Xu, T., Yang, H., & Li, S. (2022). Improving spatial visualization and Mental Rotation using FORSpatial through shapes and letters in virtual environment. *IEEE Transactions on Learning Technologies*, *15*(3), 326–337. https://doi.org/10.1109/TLT.2022.3170928.

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