Received: 20 March 2023 | Accepted: 13 June 2024

DOI: 10.1111/bjet.13499

ORIGINAL ARTICLE

British Journal of Educational Technology BERA

What factors influence scientific concept learning? A study based on the fuzzy-set qualitative comparative analysis

Jingjing Ma[1,2] | Qingtang Liu[1,2] | Shufan Yu[1,2] | Jindian Liu[3] | Xiaojuan Li[4] | Chunhua Wang[5]

[1]School of Educational Information Technology, Central China Normal University, Wuhan, China

[2]Hubei Research Center for Educational Informationization, Central China Normal University, Wuhan, China

[3]Shenzhen Donghu Middle High School, Shenzhen, China

[4]Faculty of Education, Henan Normal University, Xinxiang, China

[5]School of Animation, Huang Huai University, Zhumadian, China

Correspondence
Qingtang Liu, School of Educational Information Technology, Central China Normal University, No. 382 Xiongchu Avenue, Hongshan District, Wuhan, Hubei Province 430000, China.
Email: liuqtang@mail.ccnu.edu.cn

Funding information
Humanities and Social Sciences Research Project of the Ministry of Education of China, Grant/Award Number: 22YJAZH067; National Natural Science Foundation of China, Grant/Award Number: 72174070; Henan Higher Education Teaching Reform Research and Practice (General) Project, Grant/Award Number: 2021SJGLX531

Abstract: This research employs the fuzzy-set qualitative comparative analysis (fsQCA) method to investigate the configurations of multiple factors influencing scientific concept learning, including augmented reality (AR) technology, the concept map (CM) strategy and individual differences (eg, prior knowledge, experience and attitudes). A quasi-experiment was conducted with 194 seventh-grade students divided into four groups: AR and CM ($N=52$), AR and non-CM ($N=51$), non-AR and CM ($N=40$), non-AR and non-CM ($N=51$). These students participated in a science lesson on 'The structure of peach blossom'. This study represents students' science learning outcomes by measuring their academic performance and cognitive load. The fsQCA results reveal that: (1) factors influencing students' academic performance and cognitive load are interdependent, and a single factor cannot constitute a necessary condition for learning outcomes; (2) multiple pathways can lead to the same learning outcome, challenging the notion of a singular best path derived from traditional analysis methods; (3) the configurations of good and poor learning outcomes exhibit asymmetry. For example, high prior knowledge exists in both configurations leading to good and poor learning outcomes, depending on how other conditions are combined.

KEYWORDS
augmented reality, concept map strategy, fsQCA, individual differences, science education

Practitioner notes

What is already known about this topic

- Augmented reality proves to be a useful technological tool for improving science learning.
- The concept map can guide students to describe the relationships between concepts and make a connection between new knowledge and existing knowledge structures.
- Individual differences have been emphasized as essential external factors in controlling the effectiveness of learning.

What this paper adds

- This study innovatively employed the fsQCA analysis method to reveal the complex phenomenon of the scientific concept learning process at a fine-grained level.
- This study discussed how individual differences interact with AR and concept map strategy to influence scientific concept learning.

Implications for practice and/or policy

- No single factor present or absent is necessary for learning outcomes, but the combinations of AR and concept map strategy always obtain satisfactory learning outcomes.
- There are multiple pathways to achieving good learning outcomes rather than a single optimal solution.
- The implementation of educational interventions should fully consider students' individual differences, such as prior knowledge, experience and attitudes.

INTRODUCTION

K-12 science education is a prominent focus in the field of education, as it plays a pivotal role in nurturing students' scientific literacy and contributing to the nation's innovation capacity (Cai et al., 2022; Reiss, 2020). However, certain scientific concepts and phenomena, such as magnetic lines of induction (Liu et al., 2021) or microscopic entities like molecules and atoms (Cai et al., 2014; Liu et al., 2023), prove challenging or impossible to observe and perceive within traditional learning environments. Consequently, students often develop misconceptions and possess a fragmented, incoherent knowledge base in specific scientific areas.

Recent efforts by researchers aim to enhance science learning among students. Some studies affirm the advantages of employing augmented reality (AR) technology in enhancing students' comprehension of scientific knowledge (Sahin & Yilmaz, 2020), fostering conceptual change (Kennedy et al., 2021; Yoon et al., 2017) and cultivating inquiry skills (Kyza & Georgiou, 2019). Other research investigates the positive impacts of fine-grained processing strategies on science learning. For instance, a meta-analysis by Schroeder et al. (2018) indicates that the concept map (CM) strategy proves more effective for STEM learning compared to other instructional methods. Especially, the use of concept map strategy can effectively reduce the extraneous processing to generate and interpret the visuospatial information, thereby reducing overall cognitive load.

Nevertheless, the prevailing perspective asserts that educational media and pedagogical strategies are interconnected, mutually influencing students' learning (Clark, 1994;

Sung & Mayer, 2013). Thus, the effectiveness of AR in promoting productive science learning is believed to depend on its orchestration with pedagogical strategies (Makransky & Petersen, 2021). Additionally, individual differences may interact with the learning environment, influencing learning outcomes (Skuballa et al., 2019). For instance, students' prior knowledge (Liu et al., 2019), attitudes (Tsivitanidou et al., 2021) and technical operation experience (Chen & Wang, 2015) have been shown to account for different learning outcomes within the same learning environment.

While existing research has explored critical factors impacting science learning, little is known about how the interplay among multiple factors contributes to good or poor learning outcomes (Ling et al., 2021). Previous studies primarily utilized quantitative analysis methods (eg, *t*-test, multiple regression analysis, ANOVA) to examine the net effect of a single variable (Sahin & Yilmaz, 2020) or the interactive effect of two variables (Liu et al., 2019) on science learning, offering a single solution to explain learning outcomes. However, these analysis methods face challenges in interpreting three-way or higher-order interactions, making complex theoretical arguments difficult to test (Douglas et al., 2020).

In this study, we employ fuzzy-set qualitative comparative analysis (fsQCA) to address these research gaps. Specifically, we review critical factors proven to significantly influence scientific concept learning in existing literature. Based on complexity theory and qualitative comparative analysis (Fiss, 2007; Mason, 2008), we construct a conceptual model delineating the factors influencing student academic performance and cognitive load. This study aims to clarify how the combination of AR technology, the CM strategy and individual differences influences learning outcomes, encompassing academic performance and cognitive load. It provides a theoretical framework for comprehending the formation mechanism of scientific concept learning. Moreover, our findings provide actionable insights to enhance effective teaching and learning in K-12 science education.

LITERATURE REVIEW

Scientific concept learning and its influencing factors

Scientific concepts play a pivotal role in advancing students' science learning (National Research Council, NRC, 2012). However, various factors hinder students' performance in grasping scientific concepts, such as unobservable phenomena and technical limitations (Wu et al., 2013; Xu et al., 2022). These practical challenges impede novices from consistently acquiring and organizing knowledge systems centred on scientific concepts, leading to fragmented understanding (NRC, 2012). Therefore, to improve the quality of students' scientific concept learning, it is imperative to identify the multiple factors that significantly affect learning of scientific concepts.

Based on the Contextual Model of Learning (CML) introduced by Falk and Dierking (2000), this study sorts out the key personal and physical factors that may affect scientific concept learning. Concerning physical factors, we concentrate on technologies and strategies that effectively support scientific concept learning. As articulated in the Framework for K-12 Science Education (NRC, 2012), students are encouraged to 'use computer simulations or simulations developed with simple simulation tools as a tool for understanding and investigating aspects of a system, particularly those not readily visible to the naked eye' (p. 58). Thus, researchers suggest that AR technology can seamlessly integrate vivid virtual information with the real-world environment (Azuma, 1997), effectively conveying complex scientific information, explaining abstract concepts or demonstrating invisible phenomena in a more understandable manner (Cai et al., 2022; Liu et al., 2021; Sahin & Yilmaz, 2020; Yoon et al., 2017). Additionally, the real-time interactive feature of AR provides students with opportunities for hands-on inquiry (Yu et al., 2022), fostering a more immersive and embodied learning experience (Conley et al., 2020).

However, some studies do not support such positive effects (eg, Lai & Chang, 2021; Thees et al., 2020). As highlighted in some research (Akçayır & Akçayır, 2017; Radu, 2014; Wu et al., 2013), the reasons for AR not delivering anticipated educational effects can be attributed to: (1) the lack of necessary learning scaffolds in the AR learning process, leading students to feel confused and overwhelmed. (2) Students without prior AR experience may invest additional time and effort in operational aspects rather than focusing on the learning content.

Recognizing that AR alone may not guarantee optimal learning outcomes, it becomes imperative to integrate suitable instructional strategies into the AR learning process (Wu et al., 2013). Recent evidence supports the effectiveness of the concept map strategy (CM) in promoting students' in-depth semantic processing of learning materials (Chen et al., 2016; Chou et al., 2022; Liang et al., 2021; Novak & Cañas, 2007). In the realm of science learning, the creation or utilization of concept maps guides students in identifying essential information, discerning hierarchical relationships among scientific concepts and integrating new knowledge into their existing structures to construct a coherent system of scientific concepts (Chen et al., 2016; Schroeder et al., 2018). However, it is important to note that the concept map strategy is not universally effective (Haugwitz et al., 2010; Li et al., 2021). In a particular study, the educational effectiveness of the concept map strategy was observed only in students with low cognitive ability (Haugwitz et al., 2010). To some extent, these inconsistent findings can be attributed to the intricate interplay between instructional strategies and individuals' aptitudes (Amadieu et al., 2009).

The studies mentioned highlight the importance of considering individual differences when assessing the educational effectiveness of AR and the concept map strategy. Established research suggests that students' prior experience with AR can impact their learning performance in an AR environment (Akçayır & Akçayır, 2017; Chen & Wang, 2015). Similarly, the effectiveness of the concept map strategy is influenced by students' prior knowledge (Haugwitz et al., 2010). Additionally, a strong correlation has been identified between students' attitudes towards science and their performance in science learning (Osborne et al., 2003; Tsivitanidou et al., 2021). Therefore, this study places particular emphasis on three individual characteristics in science learning: prior knowledge, prior experience with AR and attitudes towards science.

Prior knowledge (PK) plays a crucial role in influencing students' science learning (Liu et al., 2019). Numerous studies have demonstrated that students with high prior knowledge are better equipped to apply new knowledge in problem solving compared to peers with lower prior knowledge (Chen et al., 2014). Interestingly, students with low prior knowledge can benefit from certain educational interventions, as found by Cai et al. (2014), Conley et al. (2020) and Lin et al. (2015) in the AR learning environment. Similar conclusion was observed in studies examining the effectiveness of the concept map strategy (Haugwitz et al., 2010), where students with low prior knowledge relied more on concept maps for cognitive support.

Attitudes towards science (ATS) represent a complex system of cognitive and affective dispositions influencing students' ongoing interest in scientific issues (Kind et al., 2007). Research consistently indicates a positive association between students' attitudes towards science and their achievements in science (Nuutila et al., 2018; Tsivitanidou et al., 2021). Positive attitudes correlate with successful learning outcomes, while negative attitudes may diminish the effectiveness of educational interventions (Acarli & Acarli, 2020).

Students' technology proficiency influences their self-efficacy and technology acceptance, thereby impacting their learning performance (Cázares, 2010). Specifically, students' prior experience with AR (PEAR) directly affects their learning experience in AR environments

(Chen & Wang, 2015). For example, Dunleavy et al. (2009) found that some students experienced higher cognitive load due to unfamiliarity with AR technology. However, studies on the impact of students' prior experience with AR on learning are relatively scarce, limiting definitive conclusions.

Cognitive load theory

Facilitating effective learning in a multimedia environment requires a comprehensive understanding of the learner's cognitive structure and how it interacts with the learning environment (Kirschner et al., 2011). Cognitive load theory (CLT, Sweller et al., 1998; Van Merrienboer & Sweller, 2005) provides a framework to explain how individual differences and educational interventions influence students' working memory and cognitive process. The CLT categorizes cognitive load into intrinsic, extraneous and germane cognitive load, collectively constituting a student's total cognitive load (Sweller et al., 2019). Intrinsic cognitive load (ICL) is determined by the complexity of the learning materials and the student's prior knowledge. Extraneous cognitive load (ECL) results from inappropriate instructional designs, leading to excess information processing. Germane cognitive load (GCL) involves working memory resources handling intrinsic rather than extraneous cognitive load, thus facilitating learning (Leahy & Sweller, 2016).

The positive impact of AR on learning outcomes lies in its capability to reduce students' extraneous cognitive load. AR can superimpose virtual information on real objects, preventing cognitive resource wastage and fostering information processing in the AR learning environment (Thees et al., 2020; Yu et al., 2022). For example, Thees et al. (2020) developed an AR application based on spatial and temporal contiguity principles, which could display real-time measurement data form heat conduction experiments through AR smartglasses. In contrast to traditional labs where students use a handheld thermal imaging camera to observe a metal rod and manually transmit temperature distribution still images to a computer, AR-assisted labs can reduce students' extraneous cognitive load. Nevertheless, contradictory conclusions were found in some studies (Altmeyer et al., 2020; Cheng & Tsai, 2013). For instance, Altmeyer et al. (2020) found that AR-supported lab work did not lead to less extraneous cognitive load than non-AR lab work as assumed. This research suggests that the primary factor influencing students' extraneous cognitive load is not the technology itself, but the way the AR learning environment is designed. In conclusion, there is still a lack of conclusive empirical evidence to confirm positive impact of AR on students' cognitive load, and some individual differences may affect how cognitive load is generated and handled (Ibáñez & Delgado-Kloos, 2018).

The concept map strategy affects cognitive load in two ways. On the one hand, concept maps present knowledge structures in a simpler grammatical structure, requiring fewer extraneous cognitive resources for interpretation (Schroeder et al., 2018). On the other hand, the construction of concept maps prompts students to organize learning materials into hierarchical knowledge structures, enhancing their germane cognitive load (Schroeder et al., 2018).

Students' prior knowledge decides the level of their intrinsic cognitive load (Sweller et al., 2019). Generally speaking, students with high prior knowledge retain more domain knowledge in their long-term memories and possess complete conceptual structures. These students can apply their existing cognitive structures to organize highly interactive information elements, thus reducing their intrinsic cognitive load (Ling et al., 2021). On the contrary, students with low prior knowledge usually have a higher intrinsic cognitive load, so it is possible to exceed their working memory capacity when the learning material is high in element interactivity, thereby reducing the learning efficiency (Sweller et al., 2019). Although it is commonly assumed that high prior knowledge is positively associated with good learning outcomes, some studies have found that the concept map strategy may help to prevent cognitive overload among students with lower prior knowledge, thereby resulting in more significant gains in knowledge (Haugwitz et al., 2010).

Limited studies have explored the correlation between students' learning attitudes and their cognitive load. Initial findings suggest that a positive learning attitude can result in a temporary increase in working memory capacity (Schnotz et al., 2009). According to cognitive load theory, the enhanced working memory capacity can stimulate students to process learning materials more profoundly, consequently elevating students' germane cognitive load (Bannert, 2002; Mutlu-Bayraktar et al., 2019).

Students' prior experience with AR is linked to their extraneous cognitive load. Specifically, a lack of experience in AR operations often leads to an increased extraneous cognitive load (Dunleavy et al., 2009). This is attributed to the fact that AR learning environments involve interaction procedures unrelated to the intrinsic complexity of the task, such as scanning AR markers. For students lacking AR operation experience, these interaction procedures demand more time and effort, imposing additional non-productive demands on working memory (Janssen & Kirschner, 2020).

Qualitative comparative analysis

From the aforementioned review, it is evident that science learning is a complex process, and various factors collectively impact students' academic outcomes and cognitive load. However, traditional data analysis methods (eg, *t*-test, multiple regression analysis, ANOVA) face limitations in revealing this intricate phenomenon. On the one hand, traditional quantitative analysis methods rely on symmetric tests, assuming that a predictor variable is both necessary and sufficient for the outcome (El Sawy et al., 2010; Woodside, 2013). However, the actual learning process is typically asymmetric, where AR might lead to success for some but failure for others due to individual differences. Therefore, failure cannot be simplistically viewed as the opposite of success (Ling et al., 2021; Ragin, 2008). On the other hand, traditional quantitative analysis methods overlook dependencies between multiple variables (Douglas et al., 2020). Given the complexity of the learning process and the inherent limitations of traditional quantitative analysis methods, exploring new educational data analysis approaches is crucial to unveil the dependencies among multiple factors influencing science learning.

Recently, the Qualitative Comparative Analysis (QCA) method has been introduced to the education domain (Ling et al., 2021; Nistor et al., 2019). The theoretical foundation of QCA is rooted in complexity and configuration theory, asserting that diverse variables shaping outcomes are interdependent, and the impact of a specific variable depends on its combined relationship with others, known as configurations (Fiss, 2011; Urry, 2005).

The analytical foundation of QCA is grounded in set theory, which conceptualizes conditions and outcomes as sets (Fiss, 2011; Ragin, 2008). Through the analysis of necessity and sufficiency, QCA can unveil complex causal relationship between condition sets and outcome sets (Douglas et al., 2020), such as *equifinality* (ie, multiple paths or configurations can lead to the same outcome), *conjuncture* (ie, various factors jointly influencing the outcome) and *asymmetry* (ie, causes leading to a particular outcome might be different from causes that lead to the absence of the same outcome).

QCA comes in three variants: crisp-set QCA (csQCA), multi-value QCA (mvQCA) and fuzzy-set QCA (fsQCA). fsQCA overcomes the limitations of both csQCA and mvQCA, particularly excelling in dealing with asymmetric educational data and continuous variables

(Rihoux & Ragin, 2008). This study opts for fsQCA due to its applicability to students' knowledge test scores, a type of continuous variable.

The current study

Using the fsQCA method, the main purpose of this study is to examine the necessity and sufficiency relationships between condition sets (AR, concept map strategy, prior knowledge, attitude towards science, prior experience with AR) and outcome sets (academic performance and cognitive load). Specifically, we aim to address the following three questions:

1. Is there a necessary condition that leads to good/poor academic performance and low/high cognitive load among AR, CM, prior knowledge, attitude towards science and prior experience with AR?
2. What are the specific combinations of AR, CM, prior knowledge, attitude towards science and prior experience with AR that can sufficiently cause either good or poor academic performance?
3. What are the specific combinations of AR, CM, prior knowledge, attitude towards science and prior experience with AR that can sufficiently cause either low or high cognitive load?

METHOD

The procedure of fsQCA

The procedure of fsQCA consists of the following seven steps:

1. Identify the causal conditions and outcomes. Key variables can be determined through either deductive or inductive approaches (Ketchen et al., 1993). It is advisable to avoid selecting an excessive number of conditions, as K conditions will result in 2^K possible configurations.
2. Select cases. The selected cases should demonstrate both comparability and maximum heterogeneity (Ling et al., 2021; Rihoux & Ragin, 2008).
3. Data calibration. It is necessary to convert variable values into fuzzy membership scores within the 0 to 1 range, thus establishing a fuzzy set. For binary variables, fuzzy membership scores can straightforwardly be assigned as 0 or 1. However, for continuous variables, it becomes essential to delineate three thresholds: full membership, the crossover point and full non-membership. Commonly employed calibration thresholds in the literature include 95%, 50%, 5%, or 90%, 50%, 10% (Ragin, 2008).
4. Necessary condition analysis. This step determines whether a condition is indispensable for the occurrence of a particular outcome, indicating that the outcome cannot happen without this condition. A key metric for gauging necessary conditions is consistency. Typically, the minimum accepted consistency score for identifying necessary conditions is 0.9 (Schneider & Wagemann, 2012).
5. Truth table analysis and sufficiency condition analysis. This step involves analysing whether a configuration formed by multiple antecedent conditions is a sufficient condition for an outcome (Pappas & Woodside, 2021). The initial truth table outlines all possible configurations, providing information on the frequency (ie, the number of cases for each configuration), the raw consistency and the proportional reduction in inconsistency (PRI) score (Du & Kim, 2021). It is essential to define relevant thresholds to identify configurations meeting the criteria for sufficiency conditions. For example, setting the minimum

case frequency to 1 or 2, the raw consistency value to 0.75–0.8 (Pappas et al., 2019), and the PRI scores threshold to 0.65–0.75 (Pappas & Woodside, 2021).

6. Interpret the results of the configurations with the assistance of theoretical frameworks and practical research experience.
7. Robustness test. The results of fsQCA are sensitive and stochastic due to variations in the selection of calibration point thresholds, consistency thresholds and frequency thresholds during the fsQCA analysis process. Therefore, researchers have proposed robust testing methods for fsQCA results, including fine-tuning calibration thresholds, altering case frequencies and adjusting consistency thresholds (Douglas et al., 2020; Ling et al., 2021).

The subsequent methods section primarily outlines the first two steps in the fsQCA process: the selection of condition and outcome variables, and the selection of cases. Other steps, such as data calibration, analysis of necessity and sufficiency conditions, the interpretation of fsQCA results and robustness tests, will be presented in the results section.

Identify the causal conditions and outcomes

We determined the causal conditions and outcomes based on the literature review. The conceptual model of this study is presented in a Venn diagram (Figure 1) that illustrates seven sets of constructs and their interconnections. Among them, academic performance and cognitive load are the outcome variables, while AR (Liu et al., 2021), CM (Chou et al., 2022), prior knowledge (Liu et al., 2019), attitude towards science (Kind et al., 2007) and prior experience with AR (Chen & Wang, 2015) are the causal conditions. The intersections in the diagram represent configurations of factors, which are higher-level interactions, and demonstrate the instances where one factor is present in conjunction with the others (Pappas et al., 2019).

Participants/cases selection

The study, conducted at a junior high school in central China, involved 194 seventh-grade students (aged 12–14) from four parallel classes as experimental subjects, comprising 109

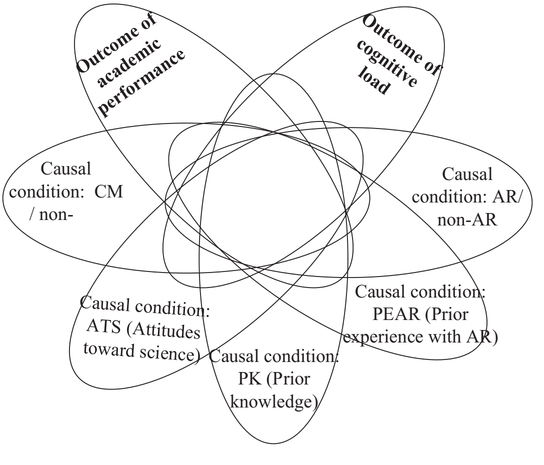

FIGURE 1 Conceptual model of scientific concept learning.

males and 85 females. The students from these four classes were divided into four groups according to whether learning with AR and CM: group1 (AR and CM, $N=52$, including 28 males and 24 females), group2 (AR and non-CM, $N=51$, including 30 males and 21 females), group3 (non-AR and CM, $N=40$, including 24 males and 16 females) and group4 (non-AR and non-CM, $N=51$, including 27 males and 24 females). Before the experiment, each participant signed an informed consent representing that they knew the purpose and process of the study. They were told their participation in the research was voluntary, and they could opt out of the experiment anytime.

For case selection, we follow the two criteria mentioned above. First, participants in this study were from four parallel classes of the same grade in the same school. These students were taught by the same biology teacher, which ensured that the selected cases were overall homogeneity. Second, in a recent biology exam, about 20% of students scored above 90, and about 10% scored below 60 (the total score was 100). It implied that large diversity exists within these cases, reflecting internal differences as much as possible and containing all possible results.

Intervention

The design of AR learning tool

The structure of plants holds significant importance in their growth and reproduction, constituting an essential topic in life science. Aligned with the Chinese junior high school biology course syllabus, this study establishes specific learning objectives: (1) comprehending the external and internal structures and functions of peach blossoms; (2) establishing hierarchical relationships between different structures and concepts; (3) understanding the pivotal role of pistil and stamen in plant reproduction.

To achieve the learning objectives and deliver the AR experience, Unity 3D software (2020.1.17) and Vuforia Engine were used to design and develop the AR learning tool, which included three learning modules:

The structure of the peach blossom

In this module (Figure 2a), students can scan the AR marker via the mobile device's camera, and the 3D model of the peach blossom will be presented on the screen. Students can rotate the AR marker freely to observe the overall external structure of the peach blossom. Furthermore, by clicking on the ‘Split’ button, students can separate the peach blossom to examine the details of each structure.

Observation of the stamen

The module (Figure 2b) displays the structure of the stamen, allowing students to use tweezers to pick up pollen and observe the internal structure of the pollen.

Observation of the pistil

This module (Figure 2c) guides students to manipulate the blade to dissect the ovary, simulating dissecting the ovary in a real-world experiment, thereby providing students with hands-on inquiry activities.

The illustration in Figure 3 portrays students in the AR group utilizing AR tools to explore the ‘The Structure of Peach Blossoms’ in the classroom. As students position the AR marker in front of the tablet's camera, a 3D model of a peach blossom appears on the screen. Students can adjust the position and angle of the 3D model by manipulating the AR marker in their hands.

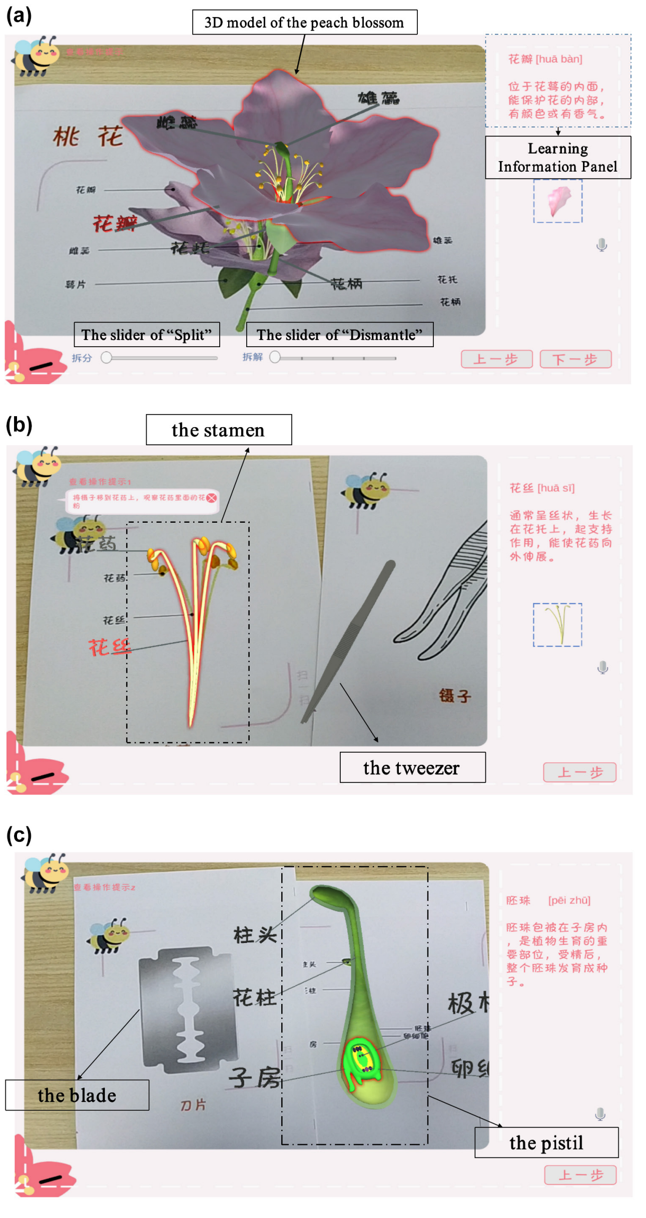

FIGURE 2 (a) The structure of the peach blossom. (b) Observation of the stamen. (c) Observation of the pistil.

For the non-AR group, they will complete the learning tasks using the textbook (Figure 4), which contains the same learning content as in the AR learning tool, but in the form of text and pictures.

The design of the concept map

The concept map stimulates students' deeper processing of learning materials. Specifically, the concept map should provide an anchoring structure that triggers students to translate and integrate new concepts into their prior knowledge networks. Therefore, this study adopted a

FIGURE 3 Students in the AR group were scanning the AR marker, and a 3D model of a peach blossom was presented on the screen.

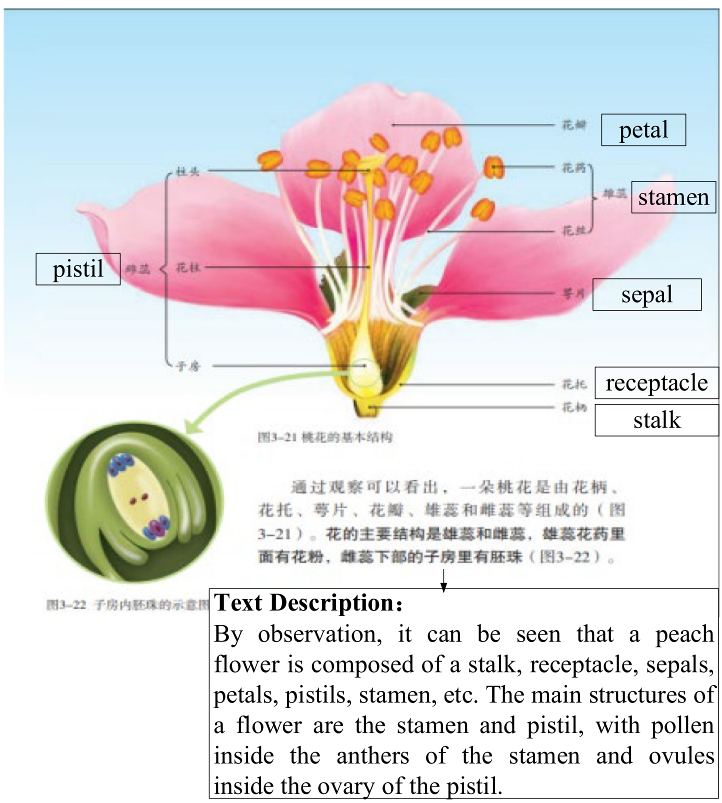

FIGURE 4 Learning materials used in the non-AR group.

filling-in-based concept map. Two biology teachers co-designed a template for the concept map (Figure 5), and students were expected to fill in the key concepts based on the clues provided by the concept map. Notably, the hierarchical structure of the concept map corresponds to the learning modules of the AR learning tool (ie, the external structure of the peach blossom—the pistil and stamen—and the internal structure of the peach blossom).

In the case of the non-CM group, they are given a worksheet (Figure 6), which does not present the hierarchical relationship between concepts like the concept map.

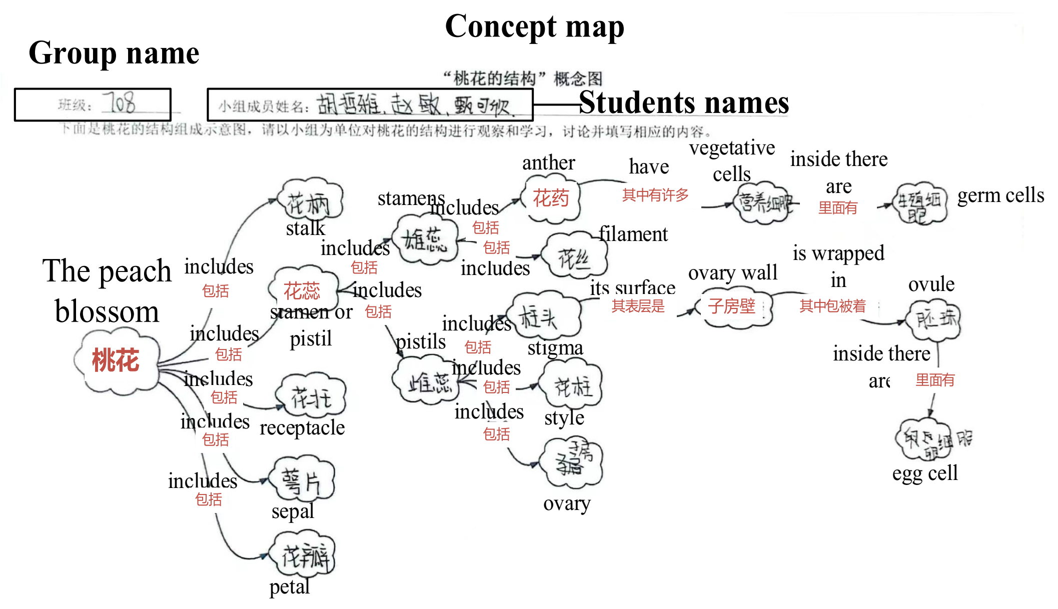

FIGURE 5 The concept map used in CM group.

FIGURE 6 The worksheet used in non-CM group.

Treatment conditions

We designed inquiry-based learning activities based on previous studies (Chiang et al., 2014; Li et al., 2010). First, the teacher created a problematic situation, asking questions such as 'What is the function of each structure of the flower? How do flowers become fruits?' to stimulate students' interest in the inquiry. Then, students worked in groups of 2–3 to conduct collaborative inquiry learning activities. Specifically, group members learned about 'the structure of peach blossoms' through the AR learning tool or the textbook (ie, non-AR) and filled out the concept map or the worksheet (ie, non-CM), with the teacher acting as a guide and technical supporter. Finally, the teacher led the class to summarize and reflect on what they had learned and started a class discussion with the question, 'Which part of the flower do you think is the most important and why?'

The following four treatment conditions were divided according to whether students learned with AR and CM.

AR and CM: In this treatment condition, each group shares a mobile device and completes three learning modules in the AR learning tool in sequence. Group members need to discuss and fill out the concept map during the learning process.
AR and non-CM: In this treatment condition, each group shares a mobile device and completes three learning modules of the AR learning tool in sequence. Group members are required to discuss and complete the worksheet.
Non-AR and CM: In this treatment condition, each group uses the textbook to learn about the structure of peach blossoms. Also, group members are required to discuss and complete the concept map.
Non-AR and non-CM: In this treatment condition, each group uses the textbook to learn the structure of peach blossoms. Also, group members need to discuss and complete the worksheet.

Measuring instruments

Pre- and post-test

The items of pre-test and post-test were co-compiled by two biology teachers with over a decade of teaching experience, which aligned with the Chinese secondary school biology course syllabus. The pre-test aimed to assess students' prior related scientific knowledge with five multiple-choice questions, one true or false question and eight fill-in-blank questions (1 point for each item, 14 points in total). The KR-20 score of pre-test (Kuder & Richardson, 1937) was 0.75, implying a reasonable internal consistency reliability. An example of a multiple choice question is 'Which part of the peach blossom can produce pollen? A: Stamens; B. Pistil: C. Petals; D. Receptacle.' An example of the true or false question is 'The stamen is not as important as the pistil because it has no relevance to the production of the fruit.' An example of a fill-in-blank question is 'In a peach blossom, the most dominant structures are _______ and _______.'

The post-test aimed to examine students' academic performance after they completed the learning tasks. The post-test adopted the same items as the pre-test and added five more multiple-choice questions (2 points for each item). The total points of the post-test were 24, and the KR-20 score was 0.72.

Cognitive load scale

The cognitive load scale was revised from the scale designed by Paas and Van Merriënboer (1993), including mental load and mental effort. The two items are 'How difficult do you think it was to learn in this lesson?' and 'How much effort do you think you devoted to the learning process in this lesson?' The scale employs the 7-point Likert rating scheme, where seven indicates high mental load and high mental effort, and one represents low mental load and low mental effort. We only adjusted the item descriptions to align with the focus of this study and did not make substantive changes to the scale. In this study, the scale's internal consistency was assessed as acceptable (Cronbach's alpha = 0.71).

Attitudes towards science scale

The scale of attitudes towards science was revised from the scale developed by Summers and Abd-El-Khalick (2018). It consists of three items with the 5-point Likert scale, where five refers to strongly agree, and one represents strongly disagree. A sample item states: 'Knowledge learned in science class is important in daily life.' The Cronbach's alpha value assessed in the current study was 0.81, indicating the scale's internal consistency was acceptable.

Prior experience with AR scale

The scale of prior experience with AR was modified from the ICT knowledge and skills subcategory of the ICT competence inventory developed by Chen and Wang (2015). It consists of five items with the 5-point Likert scale, where five refers to strongly agree, and one represents strongly disagree. A sample item states: 'I am familiar with the basic operation of AR application.' Similarly, this study did not change the substantive content of the original scale, and only customized the description of the items to AR technology. In this study, the scale's internal consistency was assessed as acceptable (Cronbach's alpha = 0.80).

Procedure

The experimental procedure includes three stages: pre-test, intervention and post-test (Figure 7). Firstly, all students completed a pre-test before the experimental intervention. The pre-test assessed students' priori relevant scientific knowledge, attitudes towards science and prior experience with AR. During the experimental intervention, four groups of students were engaged in inquiry-based learning activities about 'the structure of peach blossoms' under different learning conditions (ie, AR and CM, AR and non-CM, non-AR and

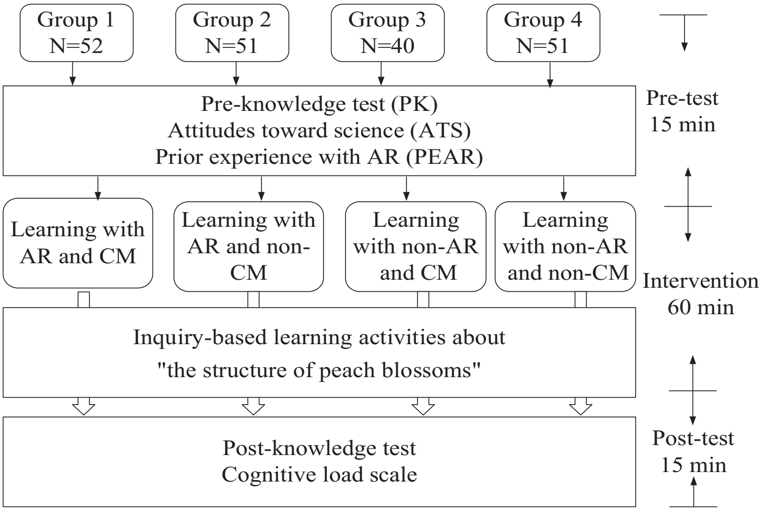

FIGURE 7 The experimental procedure.

CM, Non-AR and non-CM). Following the intervention, all students were requested to finish a post-test, including the knowledge test and the cognitive load scale. The whole experiment lasted for two class hours (ie, 90 minutes). It is worth mentioning that, to control the influence of irrelevant factors, the teaching process and experimental environment in the four experimental conditions were kept consistent. Additionally, four groups were taught by the same teacher, who has more than ten years of middle school biology teaching experience. The teacher acted as both the guide of the teaching process and the technical supporter, providing guidance to students' collaborative inquiries during the learning process.

RESULTS

Data calibration

For binary variables, we assigned 1 to CM, AR, and 0 to non-CM, non-AR. For other variables, we set the value ranked 5%, 50% and 95% as the threshold for full membership, crossover point and full non-membership (Ragin, 2008). For instance, students scoring 9 (top 5%) or higher in the pretest constituted full membership of the high prior knowledge set, while those scoring below 1.65 (bottom 5%) were considered full non-membership of the high prior knowledge set. Other students were categorized into a fuzzy set as they neither belonged to full membership nor full non-membership. Students scoring 4 (ranked 50%) were considered crossover points, indicating their maximum degree of fuzziness. Table 1 provides a summary of data calibration.

Necessary condition analysis

In this study, we employed the fsQCA 4.1 software to conduct the necessary condition analysis. It is generally accepted that if the consistency of the condition is above 0.9 (at least 0.85), this condition is considered necessary for the outcome (Schneider & Wagemann, 2012). Table 2 shows that there exists no necessary condition for the four outcomes.

TABLE 1 Data calibration.

		Anchors		
Conditions and outcomes	Mean ± *SD*	Full membership	Crossover point	Full non-membership
CM	/	1	/	0
AR	/	1	/	0
ATS	3.94 ± 0.85	5	4	2.33
PEAR	3.65 ± 0.9	5	3.8	2
PK	4.84 ± 2.32	9	4	1.65
GAP	18.22 ± 4.62	23	20	9
LCL	3.19 ± 1.4	5	2.5	1

Abbreviations: AR, augmented reality; ATS, attitudes towards science; CM, concept map strategy; GAP, good academic performance; LCL, low cognitive load; PEAR, prior experience with AR; PK, prior knowledge.

TABLE 2 Results of the necessary condition analysis.

Causal conditions	Consistency			
	GAP	PAP	LCL	HCL
CM	0.485	0.463	0.458	0.493
~CM	0.515	0.537	0.542	0.507
AR	0.591	0.467	0.424	0.656
~AR	0.409	0.533	0.576	0.344
ATS	0.681	0.641	0.679	0.635
~ATS	0.576	0.634	0.562	0.648
PEAR	0.623	0.662	0.636	0.624
~PEAR	0.659	0.641	0.613	0.669
PK	0.752	0.620	0.656	0.717
~PK	0.532	0.685	0.609	0.594

Abbreviations: AR, augmented reality; ATS, attitudes towards science; CM, concept map strategy; GAP, good academic performance; HCL, high cognitive load; LCL, low cognitive load; PAP, poor academic performance; PEAR, prior experience with AR; PK, prior knowledge; ~, the condition is absent.

Truth table analysis and the sufficiency analysis

Firstly, the fsQCA 4.1 software will output the initial truth table. Secondly, following the threshold criteria established in the existing literature (Rihoux & Ragin, 2008), the raw consistency threshold was set to 0.8, the frequency threshold was set to 2 and the PRI threshold was set to 0.75, retaining more than 80% of total cases. Finally, we obtained the results of the sufficiency analysis.

Results of fsQCA

The results of fsQCA for good academic performance and poor academic performance are shown in Table 3. First, the black circle (●) indicates the presence of this condition, the crossed-out circle (⊗) indicates the absence of this condition and the blank space indicates that the presence or absence of this condition is irrelevant to the outcome (Fiss, 2011). Second, consistency refers to the degree to which a configuration consistently results in the outcome, while overall solution consistency measures the degree to which all configurations consistently result in the outcome (Rihoux & Ragin, 2008). Generally speaking, the consistency should be larger than 0.8. Third, the raw coverage refers to the proportion of the outcome that can be attributed to a specific configuration, whereas the unique coverage pertains to the proportion of the outcome that is solely accounted for by that configuration (Rihoux & Ragin, 2008). Finally, the overall solution coverage measures the proportion of all configuration coverage in the cases of the outcome, which is comparable with the R^2 in regression analyses (Woodside, 2013).

Configurations of good academic performance

According to Table 3, the overall coverage of the three configurations (ie, A1, A2, A3) is 0.702, and the overall consistency is 0.864, indicating that 70.2% of the cases with good academic performance showed these three configurations of causal conditions (Woodside, 2013).

TABLE 3 Configurations causing good and poor academic performance.

Causal conditions	GAP			PAP		
	A1	A2	A3	B1	B2	B3
CM	●	●	●	●	⊗	⊗
AR	●	●	●	⊗	⊗	●
ATS	●		●	⊗	●	
PEAR		⊗	⊗			⊗
PK	●	●		⊗	⊗	⊗
Consistency	0.867	0.876	0.901	0.861	0.907	0.925
Raw coverage	0.413	0.366	0.315	0.514	0.482	0.349
Unique coverage	0.023	0.05	0.019	0.039	0.047	0.048
Overall consistency	0.864			0.899		
Overall coverage	0.702			0.771		

Note: A1, A2 and A3 represent the three configurations that cause good academic performance; B1, B2 and B3 represent the three configurations that cause poor academic performance. ● indicates the presence of this condition; ⊗ indicates the absence of this condition; the blank space indicates that the presence or absence of this condition is irrelevant to the outcome.

Abbreviations: AR, augmented reality; ATS, attitudes towards science; CM, concept map strategy; GAP, good academic performance; PAP, poor academic performance; PEAR, prior experience with AR; PK, prior knowledge.

First, configuration A1 shows that the combination of AR*CM*ATS*PK will achieve good academic performance. This configuration points to the top students in the class. They usually have positive attitudes towards science and solid prior knowledge, and they will get good academic performance if they can get support from both AR technology and the concept map strategy.

Second, configurations A2 (AR*CM*ATS*~PEAR) and A3 (AR*CM*PK*~PEAR) are similar. These students, despite limited AR operational experience, demonstrate positive attitudes towards science or possess high prior knowledge. These factors compensate for the lack of AR experience, enabling them to achieve satisfactory academic performance when supported by AR technology and the concept map strategy.

Configurations of poor academic performance

As shown in Table 3, three configurations (ie, B1, B2, B3) lead to PAP. The overall coverage is 0.771, and the overall consistency is 0.899, indicating that the three configurations cover 77.1% of the outcome.

First, configuration B1 (~AR*CM*~ATS*~PK) implies that students are likely to achieve poor academic performance in science learning if they lack positive attitudes towards science, do not have a strong foundation of prior knowledge and do not receive the assistance of AR, even if they are using the concept map strategy during their learning.

Second, configuration B2 indicates that if ~CM*~AR*ATS*~PK is satisfied, poor academic performance will be caused sufficiently. In this situation, even if students hold positive attitudes towards science but lack support from AR and CM in the learning environment and have low prior knowledge, they will still get poor academic performance.

Third, configuration B3 suggests that the combination of ~CM*AR*~PEAR*~PK will yield poor academic performance. Combined with Table 4, B3 (~CM*AR*~PEAR*~PK) can be regarded as a subset of D2 (~CM*AR*~PK), implying that configuration B3 will cause a high cognitive load. In this case, if these students do not have rich AR experience, it will lead to poor academic performance.

TABLE 4 Configurations causing low and high cognitive load.

Causal conditions	LCL			HCL		
	C1	C2	C3	D1	D2	D3
CM	●	⊗	●	●	⊗	⊗
AR	●	●	●	⊗	●	⊗
ATS	●		⊗			●
PEAR		●	●			
PK	●	●	⊗	⊗	⊗	●
Consistency	0.810	0.824	0.828	0.928	0.888	0.920
Raw coverage	0.342	0.312	0.238	0.484	0.363	0.521
Unique coverage	0.029	0.047	0.079	0.008	0.029	0.022
Overall consistency	0.804			0.885		
Overall coverage	0.414			0.681		

Note: C1, C2 and C3 represent the three configurations that cause low cognitive load; D1, D2 and D3 represent the three configurations that cause high cognitive load. ● indicates the presence of this condition; ⊗ indicates the absence of this condition; the blank space indicates that the presence or absence of this condition is irrelevant to the outcome.

Abbreviations: AR, augmented reality; ATS, attitudes towards science; CM, concept map strategy; LCL, low cognitive load; HCL, high cognitive load; PEAR, prior experience with AR; PK, prior knowledge.

Configurations of low cognitive load

Table 4 shows the results of fsQCA for low cognitive load. The overall coverage of 0.414 suggests that the three configurations (ie, C1, C2, C3) account for 41.4% of the cases in low cognitive load. The overall consistency (0.804) exceeds the accepted threshold of 0.80.

First, configuration C1 is the same as A1; as long as CM*AR*ATS*PK is satisfied, low cognitive load and good academic performance would be caused sufficiently at the same time, regardless of students' prior Experience with AR. This result demonstrates that students with positive attitudes towards science and solid prior knowledge will obtain low cognitive load and good academic performance if they utilize the AR learning tool to illustrate the internal and external structure of peach blossoms and get the aid of the concept map strategy to make connections between scientific concepts.

Second, configuration C2 shows that the combination of ~CM*AR*PEAR*PK can lead to low cognitive load, irrespective of students' attitude towards science. In this configuration, although students are not treated with the concept map strategy, the affordances of AR can reduce their cognitive load. Meanwhile, these students have solid prior knowledge and rich AR experience, meaning that there are more available cognitive schemas in their long-term memory which will reduce the cognitive load placed on working memory. Therefore, their total cognitive load is low.

Third, configuration C3 reveals that the combination of CM*AR*~ATS*PEAR*~PK can yield low cognitive load. Unlike configuration C1, when AR and CM are present, students with low attitudes towards science and low prior knowledge can also obtain low cognitive load, but only if such students have rich AR operation experience since it reduces the risk of increasing students' low cognitive load.

Configurations of high cognitive load

Table 4 resents three possible solutions for high cognitive load, configuration D1, D2 and D3. The overall coverage (0.681) and overall consistency (0.885) reveal that these three combinations of causal conditions covered 68.1% of the students with high cognitive load.

First, configuration D1 shows that the combination of CM*~AR*~PK will lead to high cognitive load. Notably, B1 (CM*~AR*~ATS*~PK) is a subset of D1 (CM*~AR*~PK), indicating that for students with low prior knowledge, a high cognitive load will be generated if they do not get assistance from the AR learning tool, even if they adopt the concept map strategy. In this case, if students' attitudes towards science are also not positive, it will lead to poor academic performance.

Second, configuration D2 shows that high cognitive load can be caused sufficiently when ~CM*AR*~PK is satisfied. Although these students can observe peach blossoms' internal and external structure with the aid of the AR learning tool, their low prior knowledge results in their long-term memory not storing enough schemas to process new information. In this case, without the concept map strategy to aid in organizing and assimilating new information, their working memory is more prone to being overloaded.

Third, configuration D3 (~CM*~AR*ATS*PK) indicates that for students who have solid prior knowledge and positive attitudes towards science, if they are not treated with AR and the concept map strategy, they will achieve a high cognitive load. In contrast to B2 (~CM*~AR*ATS*~PK), if the student's prior knowledge is low in this condition, it will lead to poor academic performance.

Robustness test

In this study, we conducted two methods to test the robustness of fsQCA results based on previous research (Schneider & Wagemann, 2012). Firstly, we changed the calibrating thresholds from 0.95, 0.5 and 0.05 to 0.9, 0.5 and 0.1. In other words, we changed the full membership threshold from the top 5% to the top 10% and adjusted the full non-membership from 95% to 90%. Secondly, we changed the frequency benchmark from 2 to 3, implying that the number of cases in the configuration of causal conditions included in the truth table is at least 3. After changing the parameters, we performed the necessity and sufficiency analysis again, and the results obtained were consistent with the initial results described above. Thus, the results of fsQCA in this study passed the robustness test.

DISCUSSION AND CONCLUSION

Although previous studies adopted various technologies and strategies to assist scientific concept learning, there is a lack of configurational evidence detailing these options. When past research on AR or the concept map strategy applied in the science education field is compared collectively, it reveals inconsistent empirical results, overlooks vital combinations with individual differences and obscures the true causal complexity of what drives high learning outcomes. In this study, we discuss how our research solved these research limitations. We also demonstrate how fsQCA method can be employed to analyse the complexities of the science learning process. Returning to the three research questions posed above, we draw the following conclusions.

No single factor present or absent is necessary for good/poor academic performance and high/low cognitive load

In this study, we employed the fsQCA to conduct a fine-grained analysis of cases. According to the results of the necessary condition analysis, we found that there was no single factor present or absent that is necessary for good academic performance, poor academic performance low cognitive load and high cognitive load. In detail, it was found that AR or CM alone does not always contribute to good academic performance and low cognitive load when considering individual differences (Cai et al., 2014; Haugwitz et al., 2010; Lin et al., 2015). This finding surpassed previous studies based on mean effects, which revealed a positive effect of a single factor, AR or CM, on science learning (Cai et al., 2022; Liang et al., 2021; Sahin & Yilmaz, 2020). The combinations of AR and CM always obtain satisfactory learning outcomes, such as configurations A1, A2, A3, C1 and C3. However, the prerequisite is that at least one of three individual factors (ie, positive attitude, high level of prior knowledge and AR operation experience) should be present. In other words, the antecedent conditions influencing scientific concept learning operate interdependently with each other rather than discretely or via simple two-way interactions between selected conditions. This study holds significant practical implications. Prior to the implementation of AR technology and concept map strategies, researchers or educational practitioners should proactively evaluate students' learning attitudes, technological operational experience or prior knowledge levels to ensure that students possess at least one of these prerequisites. If none of these conditions are met, researchers need to exercise caution in the utilization of AR technology and concept map strategies. In particular, pre-training can be conducted before implementing AR technology to ensure that students master basic operating skills.

Multiple, equally effective configurations of causal conditions lead to good/poor academic performance and high/low cognitive load

This study adopted a holistic perspective to reveal the complex causal relationships among AR, CM, individual differences (attitudes towards science, prior experience with AR, prior knowledge) and learning outcomes. We identified three equifinal configurations sufficient for good academic performance, poor academic performance low cognitive load and high cognitive load, respectively. The findings supported the complexity theory that there are multiple pathways to improve science learning outcomes rather than a single optimal equilibrium (Ling et al., 2021). By analysing and comparing these configurations, four main findings can be drawn:

First, it was found that AR and CM co-existing almost in all configurations of good academic performance and low cognitive load (as shown in A1, A2, A3, C1 and C2), which indicated the interdependencies of AR and CM could reduce students' cognitive load and facilitate their scientific concept learning performance. This finding reaffirms the theoretical claims of Clark (1994) and is consistent with the results of Chou et al. (2022), which revealed that combining AR and multidimensional concept maps positively influenced learning effectiveness and cognitive load. In our study, the AR learning tool superimposed the overall structure of the peach blossom and the internal structure of the pistil and stamen on the AR markers. It supplemented them with animations and audio explanations. This multimedia information based on spatiotemporal continuity would significantly reduce students' extraneous cognitive load, which has been confirmed in previous studies (Thees et al., 2020; Yu et al., 2022). Meanwhile, the hierarchical structure of the concept map used in this study corresponded to the three modules of the AR learning tool. This concept map utilized a graphical and simpler grammatical structure to help students construct the knowledge framework of the learning topic, which can effectively reduce students' extraneous cognitive load (Schroeder et al., 2018). Moreover, filling in the concept map can guide students to identify key concepts in the AR learning tool and distinguish the relationship between concepts, increasing the germane cognitive load. In summary, this finding encourages researchers to further explore the deep integration of technology and strategies to create a more effective environment for science learning.

Second, we found there were substitution effects among different variables. For instance, A2 and A3 showed that when AR and CM were present, good academic performance would be obtained if students have positive attitudes towards science or high prior knowledge. According to Skuballa et al.'s (2019) study, attitudes towards science and prior knowledge are both associated with decreased perceived task difficulty. Therefore, as long as one of these two conditions is satisfied, it can effectively reduce students' perceived task difficulty, thereby facilitating the investment of cognitive engagement. This finding also illustrates substitution effects between causal conditions, which means that the same learning outcomes can be achieved by multiple combinations of conditions (Ling et al., 2021).

Third, the analysis of the six pathways leading to poor academic performance and high cognitive load revealed that these students generally had low prior knowledge, implying a higher intrinsic cognitive load (Sweller et al., 2019). In this case, they would get a higher overall cognitive load if they could not obtain the joint assistance of AR and CM. Under such configurations, if students do not hold positive attitudes towards science or do not have high prior knowledge, it could lead to poor academic performance. This finding re-emphasized the importance of the combination of AR and CM for learning outcomes (Chen et al., 2016; Chou et al., 2022). In addition, this finding extends beyond the general viewpoint discussed in prior studies, which indicated a negative correlation between cognitive load and learning outcomes. This study further reveals that learners experiencing high cognitive load may exhibit poorer learning outcomes due to lower levels of prior knowledge or less positive learning attitudes.

Fourth, the analysis of the six pathways causing good academic performance and low cognitive load indicated that AR experience would influence cognitive load but did not directly impact learning performance. It is consistent with previous studies (Cázares, 2010; Chen & Wang, 2015; Dunleavy et al., 2009). For example, C2 and C3 showed that students should have rich experience with AR if low cognitive load is to be obtained, but A2 and A3 showed that good academic performance could be obtained even if students do not have rich experience with AR. We speculate that this may be their first time using the AR learning tool. It may generate a higher cognitive load due to unfamiliarity with how the AR system operates (eg, scanning AR markers, rotating and scaling virtual models with fingers). However, the novelty of AR keeps students highly motivated and engaged, thus allowing them to keep their attention on the learning activities. To avoid high cognitive load due to unskilled manipulation, some researchers point out that pre-training can be performed before the experiment (Dunleavy et al., 2009; Meyer et al., 2019).

There are asymmetric pathways for good and poor academic performance, high and low cognitive load

In this study, the asymmetric configurational findings provide new insights for explaining the reasons for the good and poor learning outcomes. Previous analytical approaches based on the symmetry hypothesis have validated that AR can effectively promote learning outcomes and have attributed poor performance in the control group to not benefiting from AR (Sahin & Yilmaz, 2020; Yoon et al., 2017). However, such studies ignore the role of individual differences and the inherent asymmetric relationships between the conditions and the outcomes (El Sawy et al., 2010; Woodside, 2013). We found that the configurations leading to high and low learning outcomes in the current study were not entirely symmetrical. For example, configurations C2 and D2 showed that the combination of ~CM and AR could result in both high and low cognitive load, depending on how it was combined with individual factors. This finding indicates that researchers should recognize the complexity and asymmetry inherent in the learning process. Particularly in experimental research, attributing learning failure solely to the opposite of success is overly simplistic. Instead, researchers should delve into the diverse impacts of educational interventions at the individual differences level. Future theoretical research should also incorporate complexity theory and configurational perspectives when constructing models that elucidate factors influencing learning outcomes.

IMPLICATIONS, LIMITATIONS AND FUTURE WORK

Our study theoretically and methodologically contributes to the field of K1-2 science education. This study's first theoretical implication is to understand better the complex causal relationship in scientific concept learning by integrating AR technology, the CM strategy and individual differences into a conceptual model. It is the first attempt to use fsQCA analytical method to synthesize five causal conditions to investigate their combined effects on scientific concept learning, which breaks through previous regression analysis or ANOVA studies. The second theoretical implication of this study lies in the investigation of individual differences. The role of individual differences has been widely recognized in multimedia learning research since it can be used to explain the reason for different learning outcomes in the same intervention condition. In particular, this study comprehensively investigates how individuals' priori knowledge, AR experience and attitudes influence the effectiveness of external interventions.

The present study provides some practical implications for instructional designers, instructors and relevant practitioners since it explains how critical factors of scientific concept learning predict high learning outcomes better. First, although the necessary condition analysis did not find the necessity of AR or CM alone, multiple pathways show that the combination of AR and CM could lead to satisfactory learning outcomes. Therefore, we recommend that more AR applications be introduced in K-12 science classes and science centres. Due to the abstract and microscopic nature of scientific concepts, AR technology can give full play to its technical advantages of virtual-real integration and natural interaction, superimposing virtual representational information into real situations to enhance the learning experience. More importantly, instructional designers must also select appropriate instructional strategies based on the learning content. Thus, we appeal to more studies on integrating generative learning strategies into the AR learning environment. Second, this study found that the lack of positive attitudes or high prior knowledge could be a ‘trigger’ for poor learning outcomes when students were under a high cognitive load. In this regard, teachers should pay more attention to these learners with low prior knowledge and low attitudes and provide them with individualized instruction when necessary. Third, this study found that the absence of AR experience may also yield good learning outcomes, which may benefit from the novelty effect of AR technology. Nevertheless, we suggest adding pre-training before the experiment to avoid unnecessary waste of cognitive resources caused by unskilled manipulation.

Some limitations should be identified in this study. First, in fsQCA, the number of possible configurations increases exponentially with the number of conditions (2^K), implying that an excessive number of conditions may complicate the interpretation of the findings. This study included 5 (moderate number) critical factors influencing science concept learning. Future studies can examine more complex settings involving other conditions, such as prompts or summarization strategies, VR technology and spatial ability. Second, although fsQCA is an analytical method that integrates the advantages of quantitative and qualitative analysis, it is still difficult for large-sample QCA studies to conduct qualitative analysis as deeply and richly as case studies. Therefore, in future research, we will consider interviewing students in different configurations to provide insight into the reasons for the different configurations. Third, the cognitive load measurement tool used in this study is derived from the work of

Paas and Van Merriënboer (1993). While widely employed for assessing students' cognitive load in multimedia learning environments, this tool lacks the capacity to differentiate between intrinsic, extraneous and germane cognitive loads. Hence, in future research, we will explore the adoption of the scale proposed by Klepsch et al. (2017). This scale offers a more nuanced insight into the interplay of various factors influencing students' intrinsic, extraneous and germane cognitive loads.

FUNDING INFORMATION

This work was supported by the Humanities and Social Sciences Research Project of the Ministry of Education of China [grant number 22YJAZH067]; the National Natural Science Foundation of China [grant number 72174070]; Henan Higher Education Teaching Reform Research and Practice (General) Project [grant number 2021SJGLX531].

CONFLICT OF INTEREST STATEMENT

This study has no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ETHICS STATEMENT

All procedures performed in studies involving human participants were in accordance with the ethical standards of the Research Ethics Review Committee of Faculty of Education, Central China Normal University.

ORCID

Jingjing Ma https://orcid.org/0000-0003-3507-0941
Shufan Yu https://orcid.org/0000-0001-6845-0065

REFERENCES

Acarli, D. S., & Acarli, H. A. (2020). Examination of students' attitudes towards biology and biology course in terms of gender, grade level and pet-keeping. *Problems of Education in the 21st Century*, *78*(3), 328–341.

Akçayır, M., & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review*, *20*, 1–11.

Altmeyer, K., Kapp, S., Thees, M., Malone, S., Kuhn, J., & Brünken, R. (2020). The use of augmented reality to foster conceptual knowledge acquisition in STEM laboratory courses—Theoretical background and empirical results. *British Journal of Educational Technology*, *51*(3), 611–628.

Amadieu, F., Van Gog, T., Paas, F., Tricot, A., & Mariné, C. (2009). Effects of prior knowledge and concept-map structure on disorientation, cognitive load, and learning. *Learning and Instruction*, *19*(5), 376–386.

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators & Virtual Environments*, *6*(4), 355–385.

Bannert, M. (2002). Managing cognitive load—Recent trends in cognitive load theory. *Learning and Instruction*, *12*(1), 139–146.

Cai, S., Liu, Z., Liu, C., Zhou, H., & Li, J. (2022). Effects of a BCI-based AR inquiring tool on primary students' science learning: A quasi-experimental field study. *Journal of Science Education and Technology*, *31*(6), 767–782.

Cai, S., Wang, X., & Chiang, F. K. (2014). A case study of augmented reality simulation system application in a chemistry course. *Computers in Human Behavior*, *37*, 31–40.

Cázares, A. (2010). Proficiency and attitudes toward information technologies' use in psychology undergraduates. *Computers in Human Behavior*, *26*(5), 1004–1008.

Chen, C. H., Chou, Y. Y., & Huang, C. Y. (2016). An augmented-reality-based concept map to support mobile learning for science. *The Asia-Pacific Education Researcher*, *25*, 567–578.

Chen, C. P., & Wang, C. H. (2015). Employing augmented-reality-embedded instruction to disperse the imparities of individual differences in earth science learning. *Journal of Science Education and Technology*, *24*, 835–847.

Chen, M. P., Wong, Y. T., & Wang, L. C. (2014). Effects of type of exploratory strategy and prior knowledge on middle school students' learning of chemical formulas from a 3D role-playing game. *Educational Technology Research and Development*, *62*, 163–185.

Cheng, K. H., & Tsai, C. C. (2013). Affordances of augmented reality in science learning: Suggestions for future research. *Journal of Science Education and Technology*, *22*, 449–462.

Chiang, T. H., Yang, S. J., & Hwang, G. J. (2014). Students' online interactive patterns in augmented reality-based inquiry activities. *Computers & Education*, *78*, 97–108.

Chou, Y. Y., Wu, P. F., Huang, C. Y., Chang, S. H., Huang, H. S., Lin, W. M., & Lin, M. L. (2022). Effect of digital learning using augmented reality with multidimensional concept map in elementary science course. *The Asia-Pacific Education Researcher*, *31*(4), 383–393.

Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development*, *42*, 21–29.

Conley, Q., Atkinson, R. K., Nguyen, F., & Nelson, B. C. (2020). MantarayAR: Leveraging augmented reality to teach probability and sampling. *Computers & Education*, *153*, 103895.

Douglas, E. J., Shepherd, D. A., & Prentice, C. (2020). Using fuzzy-set qualitative comparative analysis for a finer-grained understanding of entrepreneurship. *Journal of Business Venturing*, *35*(1), 105970.

Du, Y., & Kim, P. H. (2021). One size does not fit all: Strategy configurations, complex environments, and new venture performance in emerging economies. *Journal of Business Research*, *124*, 272–285.

Dunleavy, M., Dede, C., & Mitchell, R. (2009). Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of Science Education and Technology*, *18*, 7–22.

El Sawy, O. A., Malhotra, A., Park, Y., & Pavlou, P. A. (2010). Research commentary: Seeking the configurations of digital ecodynamics: It takes three to tango. *Information Systems Research*, *21*(4), 835–848.

Falk, J. H., & Dierking, L. D. (2000). *Learning from museums: Visitor experiences and the making of meaning*. AltaMira.

Fiss, P. C. (2007). A set-theoretic approach to organizational configurations. *Academy of Management Review*, *32*(4), 1180–1198.

Fiss, P. C. (2011). Building better causal theories: A fuzzy set approach to typologies in organization research. *Academy of Management Journal*, *54*(2), 393–420.

Haugwitz, M., Nesbit, J. C., & Sandmann, A. (2010). Cognitive ability and the instructional efficacy of collaborative concept mapping. *Learning and Individual Differences*, *20*(5), 536–543.

Ibáñez, M. B., & Delgado-Kloos, C. (2018). Augmented reality for STEM learning: A systematic review. *Computers & Education*, *123*, 109–123.

Janssen, J., & Kirschner, P. A. (2020). Applying collaborative cognitive load theory to computer-supported collaborative learning: Towards a research agenda. *Educational Technology Research and Development*, *68*(2), 783–805.

Kennedy, A. A., Thacker, I., Nye, B. D., Sinatra, G. M., Swartout, W., & Lindsey, E. (2021). Promoting interest, positive emotions, and knowledge using augmented reality in a museum setting. *International Journal of Science Education, Part B*, *11*(3), 242–258.

Ketchen, D. J., Jr., Thomas, J. B., & Snow, C. C. (1993). Organizational configurations and performance: A comparison of theoretical approaches. *Academy of Management Journal*, *36*(6), 1278–1313.

Kind, P., Jones, K., & Barmby, P. (2007). Developing attitudes towards science measures. *International Journal of Science Education*, *29*(7), 871–893.

Kirschner, F., Kester, L., & Corbalan, G. (2011). Cognitive load theory and multimedia learning, task characteristics and learning engagement: The current state of the art. *Computers in Human Behavior*, *27*(1), 1–4.

Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Frontiers in Psychology*, *8*, 1997.

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, *2*(3), 151–160.

Kyza, E. A., & Georgiou, Y. (2019). Scaffolding augmented reality inquiry learning: The design and investigation of the TraceReaders location-based, augmented reality platform. *Interactive Learning Environments*, *27*(2), 211–225.

Lai, J. Y., & Chang, L. T. (2021). Impacts of augmented reality apps on first graders' motivation and performance in English vocabulary learning. *SAGE Open*, *11*(4), 21582440211047549.

Leahy, W., & Sweller, J. (2016). Cognitive load theory and the effects of transient information on the modality effect. *Instructional Science*, *44*, 107–123.

Li, F. Y., Hwang, G. J., Chen, P. Y., & Lin, Y. J. (2021). Effects of a concept mapping-based two-tier test strategy on students' digital game-based learning performances and behavioral patterns. *Computers & Education*, *173*, 104293.

Li, Q., Moorman, L., & Dyjur, P. (2010). Inquiry-based learning and e-mentoring via videoconference: A study of mathematics and science learning of Canadian rural students. *Educational Technology Research and Development*, *58*, 729–753.

Liang, H. Y., Hsu, T. Y., & Hwang, G. J. (2021). Promoting children's inquiry performances in alternate reality games: A mobile concept mapping–based questioning approach. *British Journal of Educational Technology*, *52*(5), 2000–2019.

Lin, H. C. K., Chen, M. C., & Chang, C. K. (2015). Assessing the effectiveness of learning solid geometry by using an augmented reality-assisted learning system. *Interactive Learning Environments*, *23*(6), 799–810.

Ling, Y., Zhu, P., & Yu, J. (2021). Which types of learners are suitable for augmented reality? A fuzzy set analysis of learning outcomes configurations from the perspective of individual differences. *Educational Technology Research and Development*, *69*, 2985–3008.

Liu, Q., Ma, J., Yu, S., Wang, Q., & Xu, S. (2023). Effects of an augmented reality-based chemistry experiential application on student knowledge gains, learning motivation, and technology perception. *Journal of Science Education and Technology*, *32*(2), 153–167.

Liu, Q., Yu, S., Chen, W., Wang, Q., & Xu, S. (2021). The effects of an augmented reality based magnetic experimental tool on students' knowledge improvement and cognitive load. *Journal of Computer Assisted Learning*, *37*(3), 645–656.

Liu, Q. T., Liu, B. W., & Lin, Y. R. (2019). The influence of prior knowledge and collaborative online learning environment on students' argumentation in descriptive and theoretical scientific concept. *International Journal of Science Education*, *41*(2), 165–187.

Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review*, *33*, 937–958.

Mason, M. (2008). Complexity theory and the philosophy of education. *Educational Philosophy and Theory*, *40*(1), 4–18.

Meyer, O. A., Omdahl, M. K., & Makransky, G. (2019). Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment. *Computers & Education*, *140*, 103603.

Mutlu-Bayraktar, D., Cosgun, V., & Altan, T. (2019). Cognitive load in multimedia learning environments: A systematic review. *Computers & Education*, *141*, 103618.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press.

Nistor, N., Stanciu, D., Lerche, T., & Kiel, E. (2019). “I am fine with any technology, as long as it doesn't make trouble, so that I can concentrate on my study”: A case study of university students' attitude strength related to educational technology acceptance. *British Journal of Educational Technology*, *50*(5), 2557–2571.

Novak, J. D., & Cañas, A. J. (2007). Theoretical origins of concept maps, how to construct them, and uses in education. *Reflecting Education*, *3*(1), 29–42.

Nuutila, K., Tuominen, H., Tapola, A., Vainikainen, M. P., & Niemivirta, M. (2018). Consistency, longitudinal stability, and predictions of elementary school students' task interest, success expectancy, and performance in mathematics. *Learning and Instruction*, *56*, 73–83.

Osborne, J., Simon, S., Collins, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, *25*(9), 1049–1079.

Paas, F. G., & Van Merriënboer, J. J. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors*, *35*(4), 737–743.

Pappas, I. O., Giannakos, M. N., & Sampson, D. G. (2019). Fuzzy set analysis as a means to understand users of 21st-century learning systems: The case of mobile learning and reflections on learning analytics research. *Computers in Human Behavior*, *92*, 646–659.

Pappas, I. O., & Woodside, A. G. (2021). Fuzzy-set qualitative comparative analysis (fsQCA): Guidelines for research practice in information systems and marketing. *International Journal of Information Management*, *58*, 102310.

Radu, I. (2014). Augmented reality in education: A meta-review and cross-media analysis. *Personal and Ubiquitous Computing*, *18*, 1533–1543.

Ragin, C. C. (2008). *Redesigning social inquiry: Fuzzy sets and beyond*. University of Chicago Press.

Reiss, M. J. (2020). But who is it for? The history of school science in England. *Science & Education*, *29*(5), 1441–1446.

Rihoux, B., & Ragin, C. C. (2008). *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques*. Sage.

Sahin, D., & Yilmaz, R. M. (2020). The effect of augmented reality technology on middle school students' achievements and attitudes towards science education. *Computers & Education*, *144*, 103710.

Schneider, C. Q., & Wagemann, C. (2012). *Set-theoretic methods for the social sciences: A guide to qualitative comparative analysis*. Cambridge University Press.

Schnotz, W., Fries, S., & Horz, H. (2009). Motivational aspects of cognitive load theory. In *Contemporary motivation research: From global to local perspectives* (pp. 69–96). Hogrefe & Huber Publishers.

Schroeder, N. L., Nesbit, J. C., Anguiano, C. J., & Adesope, O. O. (2018). Studying and constructing concept maps: A meta-analysis. *Educational Psychology Review*, *30*, 431–455.

Skuballa, I. T., Xu, K. M., & Jarodzka, H. (2019). The impact of co-actors on cognitive load: When the mere presence of others makes learning more difficult. *Computers in Human Behavior*, *101*, 30–41.
Summers, R., & Abd-El-Khalick, F. (2018). Development and validation of an instrument to assess student attitudes toward science across grades 5 through 10. *Journal of Research in Science Teaching*, *55*(2), 172–205.
Sung, E., & Mayer, R. E. (2013). Online multimedia learning with mobile devices and desktop computers: An experimental test of Clark's methods-not-media hypothesis. *Computers in Human Behavior*, *29*(3), 639–647.
Sweller, J., van Merriënboer, J. J., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, *31*, 261–292.
Sweller, J., Van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*, 251–296.
Thees, M., Kapp, S., Strzys, M. P., Beil, F., Lukowicz, P., & Kuhn, J. (2020). Effects of augmented reality on learning and cognitive load in university physics laboratory courses. *Computers in Human Behavior*, *108*, 106316.
Tsivitanidou, O. E., Georgiou, Y., & Ioannou, A. (2021). A learning experience in inquiry-based physics with immersive virtual reality: Student perceptions and an interaction effect between conceptual gains and attitudinal profiles. *Journal of Science Education and Technology*, *30*(6), 841–861.
Urry, J. (2005). The complexity turn. *Theory, Culture & Society*, *22*(5), 1–14.
Van Merrienboer, J. J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, *17*, 147–177.
Woodside, A. G. (2013). Moving beyond multiple regression analysis to algorithms: Calling for adoption of a paradigm shift from symmetric to asymmetric thinking in data analysis and crafting theory. *Journal of Business Research*, *66*(4), 463–472.
Wu, H. K., Lee, S. W. Y., Chang, H. Y., & Liang, J. C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education*, *62*, 41–49.
Xu, W. W., Su, C. Y., Hu, Y., & Chen, C. H. (2022). Exploring the effectiveness and moderators of augmented reality on science learning: A meta-analysis. *Journal of Science Education and Technology*, *31*(5), 621–637.
Yoon, S., Anderson, E., Lin, J., & Elinich, K. (2017). How augmented reality enables conceptual understanding of challenging science content. *Journal of Educational Technology & Society*, *20*(1), 156–168.
Yu, S., Liu, Q., Ma, J., Le, H., & Ba, S. (2022). Applying augmented reality to enhance physics laboratory experience: Does learning anxiety matter? *Interactive Learning Environments*, *31*, 1–16.

How to cite this article: Ma, J., Liu, Q., Yu, S., Liu, J., Li, X., & Wang, C. (2024). What factors influence scientific concept learning? A study based on the fuzzy-set qualitative comparative analysis. *British Journal of Educational Technology*, *00*, 1–26. https://doi.org/10.1111/bjet.13499