PHYSICAL REVIEW PHYSICS EDUCATION RESEARCH **20,** 010112 (2024)

Analysis of visual-based physics questions of the senior high school entrance examination in China

Chengran Wang and Bing Wei[*]

Faculty of Education, University of Macau, Taipa, 999078, Macau, China

(Received 26 September 2023; accepted 12 February 2024; published 5 March 2024)

The notion of scientific visual literacy has been advocated in recent science curriculum reform documents and related learning outcomes are expected from students. However, few studies have been conducted to determine how it is tested in high-stakes examinations. This study utilized the Visualization Blooming Tool to examine the level of visual cognition involved in visual-based physics questions in the Senior High School Entrance Examination (SHSEE) in China. Content analysis was adopted as the research method and 12 sets of the SHSEE physics from four Chinese metropolises (Beijing, Shanghai, Hangzhou, and Suzhou) in 2020, 2021, and 2022 were targeted. The results indicate that although all four metropolises examined the higher-order visual cognitive skills, they placed more emphasis on the levels of apply and analyze but less on evaluate and create. Moreover, the examination items required students to interpret visual representations more often than to construct them, which may be detrimental to developing students' scientific visual literacy. It is suggested that the examination of higher-order visual cognitive skills and encoded visual representation should be strengthened in future high-stakes physics examinations.

DOI: 10.1103/PhysRevPhysEducRes.20.010112

I. INTRODUCTION

For over a decade, scientific visual literacy has been increasingly regarded as a prerequisite for scientific communication and has gained considerable attention in science education research [1–3]. It has recently been advocated in curriculum reform documents around the world and demanding outcomes are expected from students [4–6]. For instance, in the latest version of the national physics curriculum standards of junior high schools (grades 7–9) in China, it is explicitly stated that after completing physics studies in the stage of compulsory education, students should be able to understand the meaning of visual representations and construct visual representations to solve complex problems in real-life situations [5]. To achieve success in standards-based curriculum reforms, students' scientific visual literacy should be covered in the examination. However, so far, there has not been sufficient information about how and to what extent it is tested in external examinations, especially in high-stakes examinations. As far as we know, three studies have explored the incorporation of visual representations in high-stakes examinations but all of them focused on the types of visual representation employed [7–9]. For example, LaDue *et al.* [7] examined the 2012 New York State Regents examination and discovered that all four science subjects (earth science, biology, chemistry, and physics) included graphs, tables, and diagrams in the examination items, reflecting the prevalence of the visual representations recommended in the curriculum documents. The current study aims to investigate to what extent scientific visual literacy is included in high-stakes examinations from the perspective of students' cognitive processing.

The Senior High School Entrance Examination (SHSEE), commonly known as *Zhongkao*, is an important high-stakes examination that has been widely valued by the Chinese society [10]. For junior high school graduates (approximately 15 years old), the SHSEE serves a multifaceted role. Not only does it assess their academic achievements, determine their graduation, and influence further education pursuits but also serves as a screening tool with examination results often being considered the only referent for admission to senior high schools. Adding to its significance, the SHSEE is conducted only once annually, further intensifying its high-stakes nature. Physics is a compulsory subject in junior high schools in China, as well as one of the subjects in the SHSEE. As a content-based examination, the SHSEE physics attaches importance to the examination of visual cognitive skills, which is considered to be an important way to evaluate students' scientific visual literacy. For instance, in the 2022 Beijing SHSEE physics, there are 42 questions, about 80% of which involve measuring visual skills such as the interpretation of diagrams, photos, and graphs, and the use of diagrams to

---

[*] Corresponding author: bingwei@um.edu.mo

describe information. In this sense, it is necessary to examine the extent to which different visual cognitive skills underpinning scientific visual literacy are tested by the SHSEE physics and thus determine whether the examinations echo the advocacy made in the intended curriculum for visual cognitive skills, especially higher-order cognitive skills.

The current study is based on two considerations: First, research has shown that students are motivated to perform well on high-stakes examinations; thus, the cognitive challenge of examination items has a profound impact on their learning strategies [11]. If the examination items highlight lower-order visual cognitive skills, students are likely to focus more on the visual representation itself, rather than connecting the visual representation with conceptual interpretation, which may hinder the development of scientific visual literacy [12]. Therefore, it is necessary for the examination items to be oriented toward higher-order visual cognitive skills. Second, interpreting (e.g., multiple choice) and constructing (e.g., free graphical response) visual representations are two tasks that involve the bidirectional translation of text and visual [8]. As multiple choice possesses fixed answers and covers a wide range of topics [13], they are frequently used to assess factual knowledge, whereas the examination of higher-order cognitive skills, particularly the category of create, is limited [14]. Free graphical responses, on the other hand, require students to extract, analyze, apply, and evaluate the printed information they read, and finally construct their own visual representations to solve problems [15,16]. Due to the task complexity, free graphical responses often require students to employ higher-order visual cognitive skills compared to multiple-choice questions [16,17]. Given that constructing visual representations is beneficial for students' in-depth learning of science [18], it is worthwhile to examine whether they are adequately utilized in high-stakes examinations.

Departing from the above considerations, this study took 12 sets of the SHSEE physics from four Chinese metropolises (Beijing, Shanghai, Hangzhou, and Suzhou) in 2020, 2021, and 2022 as the analysis target and used the Visualization Blooming Tool (VBT) [19] to examine visual cognitive skills of visual-based questions involved in these examinations. The research questions are raised as follows:

(1) How are visual cognitive skills distributed in the 2020–2022 SHSEE physics among four Chinese metropolises?
(2) How are visual cognitive skills distributed in different answer formats in the 2020–2022 SHSEE physics among four Chinese metropolises?

II. LITERATURE REVIEW

A. Visual representation and visual literacy

Visual representation is one of the essential components of science and plays an indispensable role in the process of science communication [20]. Compared with verbal representation and symbolic representation, visual representation is vivid and intuitive. This is why visual representation is often used to describe abstract concepts, which is conducive to the understanding of scientific concepts and the construction of scientific laws [8]. Additionally, visual representation can be viewed as an effective tool to convey complex information, allowing scientists to make hypotheses, identify meaningful patterns in data, and communicate their ideas to the general public and other scientists [19]. In the field of K-12 science education, visual representation is widely used in textbooks and teachers' instruction to disseminate scientific knowledge in a more effective manner [21,22]. However, it has been evidenced that science teachers have difficulty conveying information contained within visual representations to their students [23]. This is because students need to possess not only relevant content knowledge but also sufficient visual literacy [20]. In recent decades, visual literacy has received considerable attention in a variety of fields, particularly in the field of education [24–28]. Some education scholars proposed that, like verbal literacy, visual literacy is a communication skill that can be defined as the ability to interpret and construct visual representations [26–28]. It should be noted that visual literacy is not a generic skill but rather a discipline-specific one [20,29]. In other words, the same visual convention may have different meanings across disciplines. For example, a line segment with an arrow can represent the sequence of events over time in history, while in physics, it can be used to represent the magnitude and direction of forces. Only by learning and becoming familiar with the visual conventions of the specific disciplinary ways of knowing, will students attain disciplinary discourse fluency, and thus facilitate the mastery of disciplinary knowledge [30]. According to Offerdahl *et al.* [31], the acquisition of visual literacy in a discipline is a process of achieving disciplinary discourse fluency, therefore, scientific visual literacy can be defined as "the achievement of fluency in the disciplinary discourse scientists use when engaging in activities such as (1) decoding and interpreting visual representations, (2) encoding and creating visual representations, and (3) generating mental models" (p. 2). These hierarchical visualization skills underpinning scientific visual literacy reflect different cognitive levels, which can be analyzed by Bloom's taxonomy [11,19].

B. Bloom's taxonomy and assessment of visualization skills

Bloom's taxonomy has received considerable attention from scholars in education, particularly in science education [32–35]. According to Bloom's taxonomy, learning objectives can be classified into three domains: cognitive, affective, and psychomotor. As for the cognitive domain, it consists of six different categories, including knowledge, comprehension, application, analysis, synthesis, and evaluation [36]. In response to changes in the educational environment and the development of curriculum theory, Anderson *et al.* [37] revised Bloom's original taxonomy. To facilitate operation and application, the revised version expands the structure from one to two dimensions, "Knowledge" and "Cognitive Process." In addition, the terminology of the cognitive process dimension has been modified from nouns to verbs, which includes six categories: remember, understand, apply, analyze, evaluate, and create [37].

As indicated in the literature, Bloom's taxonomy can be applied to analyze a wide range of literacies, such as mathematical literacy, scientific literacy, and visual literacy [38,39]. For scientific visual literacy in particular, Trumbo [20] used Bloom's taxonomy to describe visual learning strategies in science communication, including familiarity with specific visual representations of scientific disciplines and interpretation of meaning from visual representations. Other scholars have designed visual tasks based on revised Bloom's taxonomy to develop students' scientific visual literacy [11,19,38]. For example, Crowe *et al.* [11] developed the Blooming Biology Tool based on revised Bloom's taxonomy, which involves specific descriptions of each level of visual cognition, assisting biology teachers in designing examination items for assessing students' mastery of visual cognitive skills and to help them improve higher-order visual cognitive skills. Mnguni *et al.* [38] used revised Bloom's Taxonomy to identify the visual cognitive skills for optimal scientific visual literacy. As part of their study, they designed multiple questions for each visual cognitive skill based on specific biochemical concepts, requiring students to search for answers based on their visual cognitive skills and conceptual knowledge.

In these studies of designing test questions related to scientific visual literacy, Arneson and Offerdahl's [19] work deserves particular attention, as they developed the Visualization Blooming Tool (VBT), which provides a more systematic framework for designing examination items to enhance specific visual cognitive skills. In their study, the VBT was used to generate new practice and examination items to reinforce their alignment in visual cognitive skills in the biochemistry course. Apart from generating new questions, the VBT could also be used to examine the existing examination items, aiming to evaluate the extent to which different visual cognitive skills are assessed in these items [19]. However, few research has focused on the assessment of visual cognitive skills in existing examinations, especially in high-stakes examinations. To fill this gap, this study used the VBT to categorize visual-based physics questions in the SHSEE based on their visual cognitive skills.

C. Visualization blooming tool

As stated earlier, the Visualization Blooming Tool (VBT) is deemed a useful tool for categorizing existing examination items based on their visual cognitive skills. In addition, research has shown that the VBT is applicable to a wide variety of scientific disciplines, such as biology, physics, and chemistry [19]. We believed it would be appropriate to analyze the visual-based physics questions of the SHSEE in the present study.

In accordance with revised Bloom's taxonomy of the cognitive domain, the VBT classifies visual cognitive skills into six categories [37]. Among them, *remember* and *understand* are considered lower-order cognitive skills

TABLE I. The Visualization Blooming Tool (adapted from Ref. [19]). Note that LOCS indicates lower-order cognitive skills; HOCS indicates higher-order cognitive skills.

Categories	Cognitive processes	General characteristics	Examples of visual tasks
Remember (LOCS)	Recognize Identify Recall Retrieve	Students only need to recall the relevant facts or information to answer the items correctly.	• Recognize experimental process or conventional method that would yield the representation. • List ordered steps in a schematic. • Identify the structure or characteristics of an object. • Define abbreviations or symbols in the representation.
Understand (LOCS)	Interpret Exemplify Classify Summarize Infer Compare Explain	The context of the items is usually familiar to students, and they only need to focus on the surface features or the representation itself and construct meaning accordingly.	• Select a specific representation based on a defined feature of the general concept. • Categorize the representation based on their surface features. • Briefly summarize the information in the representation. • Find a pattern within a series of representations. • Detect similarities or differences between two or more representations. • Make predictions about situations that have already been explicitly covered.

(Table continued)

TABLE I. *(Continued)*

Categories	Cognitive processes	General characteristics	Examples of visual tasks
Apply (LOCS/ HOCS)	Execute Implement Use	Items require students to apply a process or procedure to solve a problem. LOCS (execute) items usually involve familiar tasks, thus providing sufficient clues for students to choose the appropriate procedure to use. Moreover, such items consist of a sequence of steps that generally follow a fixed order, and the answers are predetermined. HOCS (implement) items require conceptual understanding, situations that are usually unfamiliar to students, and require selection, modification, or production of procedures. These items may contain decision points and have no single, fixed answer.	• Calculate a solution. • Read the value of the measuring instrument. • Sketch graph from provided data. • Draw a representation depicting the expected outcome. • Predict the impact of a change in a single variable on representation. • Transform the information from one form to another.
Analyze (HOCS)	Differentiate Organize Discriminate Attribute	Items require students to understand not only the representation itself but also some background knowledge. Students need to discriminate the relevant pieces of messages, determine how each part fits into the overall structure, build connections, or underly the purpose of the representation.	• Distinguish relevant and irrelevant information in a presentation based on content knowledge. • Infer the physics meaning of the data from the representation. • Predict how representation would change if multiple properties were changed. • Make determination regarding a concept by comparing representations. • Determine the purpose for presenting the representation.
Evaluate (HOCS)	Detect Check Critique Judge	Items require students to make judgments based on criteria or standards (including quality, effectiveness, efficiency, or consistency). Students are expected to examine products or representations for internal consistency or to make judgments regarding their positive and negative characteristics based on external criteria.	• Check whether the data support or disconfirm the hypothesis. • Judge whether the representation contains parts that contradict one another. • Discern the most effective solution to the problem. • Determine which convention should be used in the representation to convey information. • Evaluate the effectiveness of a representation. • Critique existing representations based on the principles of physics.
Create (HOCS)	Generate Hypothesize Plan Design Produce Invent	Items require students to put elements together to form a functional whole or reorganize some elements into a pattern or structure not clearly present before. Students are expected to try to understand the task and generate a variety of possible solutions, modify the solution method, turn it into a plan of action, execute the plan, and construct the solution.	• Generate hypothesis from existing representations. • Structure evidence into an argument for or against a conclusion or hypothesis. • Design a plan to collect evidence to support scientific argument. • Develop a plan for solving the problem by using appropriate equations, variables, etc. • Generate alternative ways to represent data or information. • Provide a new representation that meets the requirements of the item.

(LOCS); *analyze*, *evaluate*, and *create* are higher-order cognitive skills (HOCS). The apply level involves two situations: if the question merely requires procedural knowledge to answer and the context is familiar to the student, such a question is considered LOCS; whereas if the question requires not only procedural knowledge but also relevant conceptual knowledge to answer and the context is unfamiliar, it is considered HOCS [19]. In addition, we modified some of the descriptions of visual tasks in the VBT to match the actual situation in the SHSEE physics. For instance, we omitted the description of "label components of the image" because this task was not found in the SHSEE physics. Another example involves requiring students to identify the value of an instrument (e.g., an ammeter), these visual tasks occur frequently in the SHSEE physics, so we included them on the VBT and classified them as apply level (see Table I).

III. RESEARCH METHOD

Content analysis is a systematic, replicable research technique for compressing many words of text (or other meaningful matter) into fewer content categories based on explicit rules of coding [40]. According to Stemler [41], content analysis can be used with a wide variety of data sources, including textual data, visual stimuli (e.g., photographs or videos), and audio data. In this study, this approach is considered suitable for examining visual-based physics questions in the SHSEE by coding visual-based questions and converting them into quantitative data such as frequency and percentages, allowing researchers to answer the two research questions raised earlier.

A. Analysis target

As indicated by the results of the Program for International Student Assessment (PISA) 2018, students from four Chinese provinces or municipalities, i.e., Beijing, Shanghai, Jiangsu, and Zhejiang, performed exceptionally well, achieving the highest marks for various literacy assessments, including scientific literacy [42]. In fact, the four provinces or municipalities are relatively economically developed in China with well-established educational assessment systems, which have greatly contributed to the reform of the Chinese education system. In terms of the SHSEE physics, different provinces or municipalities use different approaches to designing the examinations. To be more specific, in Beijing and Shanghai, they are designed directly by the municipalities, while in Zhejiang and Jiangsu by individual prefecture-level metropolises within the provinces. In this study, we selected Hangzhou and Suzhou, two prefecture-level metropolises with relatively developed education in Zhejiang and Jiangsu provinces, as representatives of the two provinces, respectively. Finally, it was determined to analyze 12 sets of the SHSEE physics involved with visual-based questions from four Chinese metropolises (Beijing, Shanghai, Hangzhou, and Suzhou) over the past 3 years (2020–2022). The printed versions of the SHSEE physics questions are publicly accessible and no ethical issue is involved.

TABLE II. The number of analysis units in the SHSEE physics of four metropolises (2020–2022).

	2020	2021	2022	Total
Beijing	37	27	34	98
Shanghai	25	15	18	58
Hangzhou	23	19	17	59
Suzhou	51	52	46	149
Total	136	113	115	364

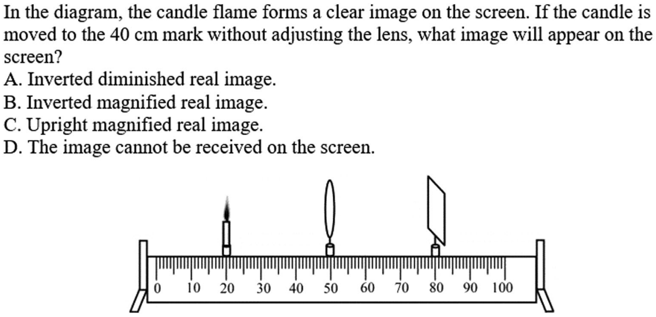
B. Analysis units

In our study, so-called "visual-based questions" included two types: one is that all or part of the known conditions of the main question are presented by visual representation; the other is that the main question itself does not provide visual representations but requires students to generate their own visual representations to answer it. The unit of analysis plays a vital role in the content analysis process [40]. In the four metropolises' SHSEE physics, a main question sometimes consisted of several visual tasks. Although these visual tasks are associated with the same knowledge unit, they may be contextually independent and address different levels of visual cognition. Therefore, each visual task was considered a unit of analysis. For example, "Please draw the lever's resistance arm and the minimum force required to maintain the lever in its balanced position in the diagram" (2022 Suzhou SHSEE physics, Question 22). This item involves two visual tasks: drawing the lever's resistance arm (*apply LOCS*) and determining the direction of the minimum force required to balance the lever (*apply HOCS*). In total, 364 units of analysis were finally extracted from the SHSEE physics of the four metropolises during the period 2020–2022 (see Table II).

C. Analysis process

First of all, the analysis units of physics questions in the SHSEE of 2020, 2021, and 2022 in Beijing, Shanghai, Hangzhou, and Suzhou were arranged in chronological order and recorded in an Excel spreadsheet. Each unit of analysis was coded as a row in the spreadsheet, including year, city, item sequence, visual cognitive skill, and response format. Based on the spreadsheet, content analysis was conducted with regard to two aspects to address the two research questions of this study, respectively.

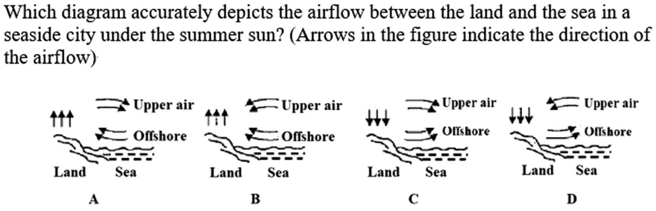
Regarding visual cognitive skills, we referred to the description of visual tasks in Table I to determine which of the seven categories a certain visual task fell into. In this process, in most cases, it was easy to classify the visuals by using the criteria of VBT. For example, some visual



In the diagram, the candle flame forms a clear image on the screen. If the candle is moved to the 40 cm mark without adjusting the lens, what image will appear on the screen?

A. Inverted diminished real image.
B. Inverted magnified real image.
C. Upright magnified real image.
D. The image cannot be received on the screen.

FIG. 1. Sample item coded in verbal or numerical forms (B1): From the 2020 Suzhou SHSEE Physics, Item 7.



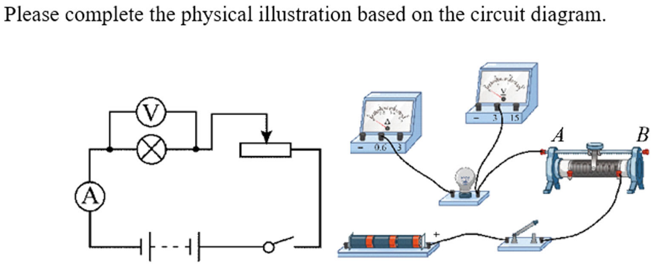
Which diagram accurately depicts the airflow between the land and the sea in a seaside city under the summer sun? (Arrows in the figure indicate the direction of the airflow)

FIG. 3. Sample item coded in visual cues (B3): From the 2020 Hangzhou SHSEE Physics, Item 6.

representations simply repeated the text without containing any new information (e.g., just visualizing the parachute mentioned in the text). In these cases, the questions were coded as remember. But in some cases, coding visual representations was challenging, especially in distinguishing between apply LOCS and apply HOCS. This required us to code the questions in combination with the situation. For example, "Please draw the moment arm of the lever and calculate the value of the force." The task only presented a diagram that the students were familiar with, thus the code was apply LOCS. In contrast, when the lever was turned, it was required to draw the moment arm corresponding to the minimum force and calculate the value of the minimum force; students had to make decisions and solve the question in an unfamiliar situation, so the task was coded as apply HOCS. To eliminate the familiarity bias, we also broke each visual task down and listed relevant knowledge and visual skills students need to complete it. For instance, the task showed the graph of temperature change over time when water is boiling. If the task required students to identify the value of boiling point, which means students only requested to make a judgment based on the visual representation itself, it was coded as the understand level. However, if the task asked students to judge the local pressure situation based on the graph, students not only had to recognize the boiling point value from it, but also relate it to their conceptual knowledge of atmospheric pressure, so it was coded as the analyze level. Further, a visual-based question might involve multiple visual tasks; in this case, it was necessary to take into consideration the relationship between them. For example, a question contained two tasks, the first of which required students to identify a circuit fault by selecting the most appropriate solution from several options provided (evaluate level); and the second one required students to determine how the ammeter and voltage indications changed based on the scheme selected in the first task. If we looked at the second task separately, it was coded at the analyze level. However, since students had already gone through the process of analysis when they answered the first task, the second task was merely to organize the language to express possible phenomena based on the findings of the first task, thus it was categorized as the understand level.

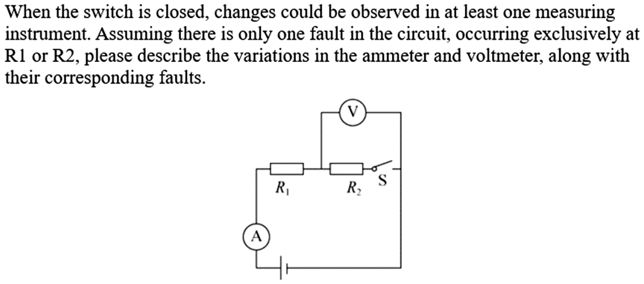
With regard to response formats, we followed Yeh and McTigue's [8] approach and coded the visual tasks into two categories: built-in answers and free responses. Built-in answers refer to items with choice answers embedded either in the posed question or answer options, including (i) verbal or numerical forms (B1) (with multiple choices), (ii) graphics completion (B2), e.g., complete the physical illustration based on the existing circuit diagram, and (iii) visual cues (B3), e.g., select the correct graphical object. Free responses have open-ended formats, including (i) free verbal or numerical response (F1), e.g., describe an experimental phenomenon, (ii) free graphical response (F2), e.g., draw a force analysis diagram, and (iii) chart completion with free verbal response (F3), e.g., fill in the data in the table (see Figs. 1–6).

To ensure the objectivity of the content analysis, two raters independently encoded 364 analysis units using the VBT tool. The two raters, a postgraduate student



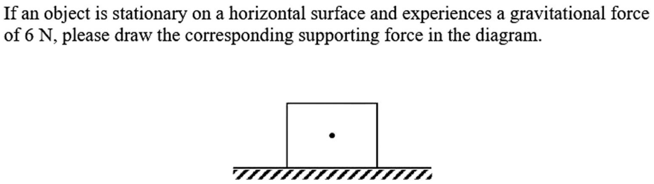
Please complete the physical illustration based on the circuit diagram.

FIG. 2. Sample item coded in graphics completion (B2): From the 2021 Hangzhou SHSEE Physics, Item 10.



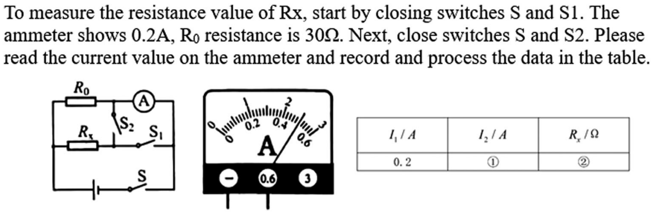
When the switch is closed, changes could be observed in at least one measuring instrument. Assuming there is only one fault in the circuit, occurring exclusively at R1 or R2, please describe the variations in the ammeter and voltmeter, along with their corresponding faults.

FIG. 4. Sample item coded in free verbal or numerical response (F1): From the 2021 Shanghai SHSEE Physics, Item 12.



If an object is stationary on a horizontal surface and experiences a gravitational force of 6 N, please draw the corresponding supporting force in the diagram.

FIG. 5. Sample item coded in free graphical response (F2): From the 2022 Shanghai SHSEE Physics, Item 14.



To measure the resistance value of Rx, start by closing switches S and S1. The ammeter shows 0.2A, R_0 resistance is 30Ω. Next, close switches S and S2. Please read the current value on the ammeter and record and process the data in the table.

I_1/A	I_2/A	R_x/Ω
0.2	①	②

FIG. 6. Sample item coded in chart completion with a free verbal response (F3): From the 2020 Beijing SHSEE Physics, Item 30. Both visual tasks (① and ②) were coded as F3.

(former physics teacher) and a university professor in science education, have experience in interpreting curriculum standards, teaching content, and Bloom's taxonomy applications, which significantly contributed to the completion of this study. To ensure the reliability of the study, we first compared the results of encoding visual cognitive skills by the two raters, the percentage of agreement was 85%. Cohen's kappa, which is used to measure the degree of agreement between two raters after excluding the possibility of chance coding [43], was also calculated ($\kappa = 0.814$). According to Blackman and Koval [44], a kappa value greater than or equal to 0.80 is considered a nearly perfect agreement. In cases of disagreement, most of which resulted from different understandings of the cognitive processes of specific items, the two raters discussed their differences based on the VBT and finally reached an agreement.

IV. RESULTS

In this section, the results of data analysis are presented under two subsections to, respectively, answer the two research questions raised earlier.

A. The distribution of visual cognitive skills in the SHSEE physics among four metropolises

The number of analysis units in the SHSEE physics among the four metropolises varied greatly. For instance, only 15 analysis units were contained in Shanghai 2021, while up to 52 analysis units were in Suzhou 2021. For the purpose of the study, the distribution of visual cognitive skills among the four metropolises is presented with the percentages (see Fig. 7).

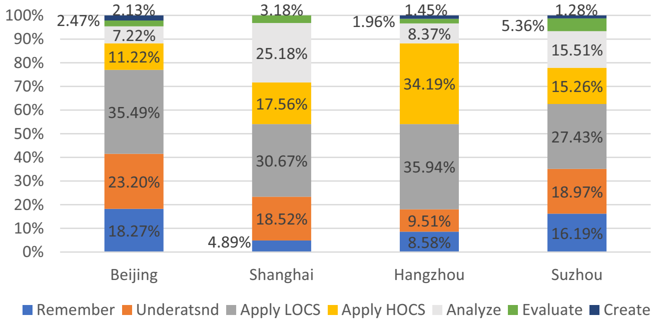


FIG. 7. The distribution of visual cognitive skills in the SHSEE physics among four metropolises.

As shown in Fig. 7, the SHSEE physics of the four metropolises all covered both HOCS and LOCS visual tasks, but each placed a different emphasis on the level of visual cognition. Specifically, Beijing had the highest percentage of LOCS visual tasks (76.96%) (including remember, understand, and *apply* LOCS), while HOCS visual tasks (namely, apply HOCS, analyze, evaluate, and create) accounted for the lowest percentage (23.04%). That is to say, Beijing had relatively low requirements for visual cognitive skills, and most of the visual tasks only required students to understand the visual representation itself. Compared to Beijing, Suzhou had a slightly lower percentage of LOCS visual tasks (62.59%), and the percentage of HOCS visual tasks was 37.41%. Unlike Beijing and Suzhou, which focused on LOCS visual tasks, Shanghai and Hangzhou examined both HOCS and LOCS visual tasks in a relatively balanced manner. LOCS visual tasks accounted for 54.08% and 54.03% in Shanghai and Hangzhou, respectively, and the HOCS visual tasks accounted for 45.92% and 45.97% in Shanghai and Hangzhou, respectively, demonstrating higher expectations for student's scientific visual literacy. However, it should be noted that Shanghai and Hangzhou put different emphases on specific HOCS visual tasks. Shanghai placed a greater emphasis on the level of analyze (25.18%), and students were frequently asked to extract useful information from visual representations to answer some complex questions. Hangzhou, on the other hand, emphasized the level of apply (34.19%), that is, students were often required to apply relevant conceptual knowledge to implement nonspecific procedures in unfamiliar situations.

By analyzing the specific visual cognitive skills involved in the visual-based questions, we found that apply accounted for the highest percentage of LOCS visual tasks. This may be because students were required to complete basic visual tasks, such as reading the value of a measuring instrument or calculating a solution. In addition, for HOCS visual tasks, the four metropolises generally emphasized the apply and analyze, while evaluate and create were neglected. It is particularly noteworthy that none of the visual-based questions from Shanghai examined the create over the past 3 years.

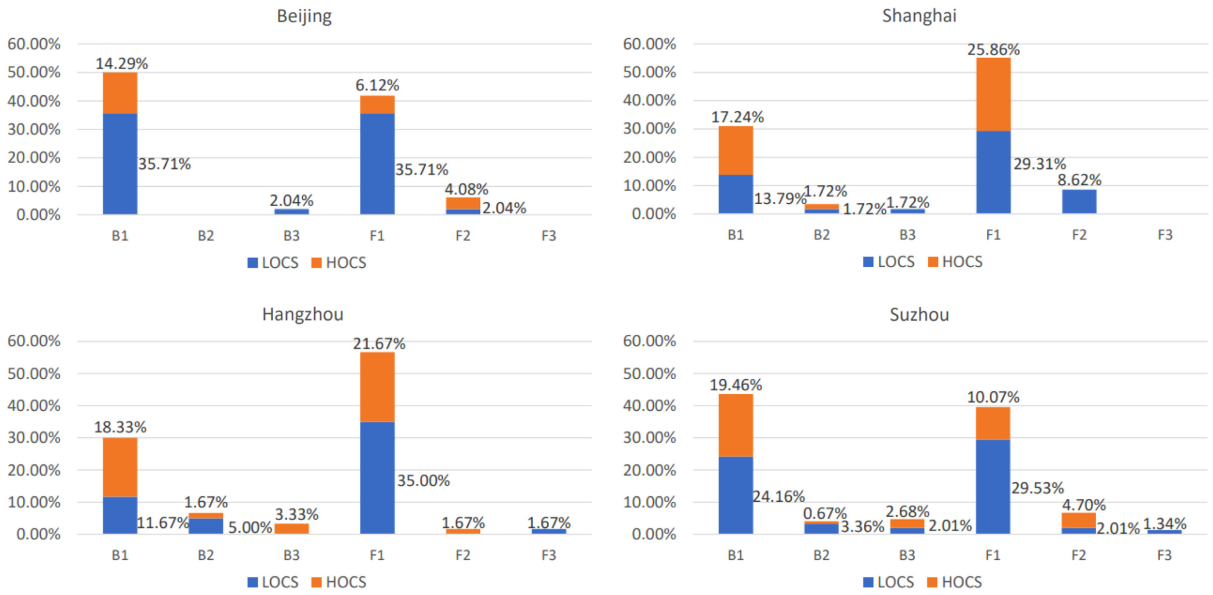


FIG. 8. The distribution of visual cognitive skills in different answer formats in the SHSEE physics among four metropolises.

B. The distribution of visual cognitive skills in different answer formats in the SHSEE physics among four metropolises

With regard to the second research question, we examined the response formats for the visual-based questions in the SHSEE physics of the four metropolises and calculated the average percentages of different response formats, to determine whether the different response formats reflect different levels of visual cognition (see Fig. 8).

As can be seen from Fig. 8, there are relatively few built-in answers (the sum of B1, B2, and B3) in Shanghai (36.19%) and Hangzhou (40.00%). In contrast, Beijing and Suzhou have relatively high percentages of built-in answers (the sum of B1, B2, and B3), accounting for 52.04% and 52.34%, respectively. In this study, multiple choice forms were coded into verbal or numeric forms (B1). Beijing and Suzhou have a high proportion of multiple choices, which may lead to a high number of built-in answers. According to Frederiksen, *the difficulty of devising multiple-choice items that measure higher-order cognitive skills tends to result in tests that elicit factual knowledge rather than more complex cognitive processes* [13]. However, in the SHSEE physics of the four metropolises, this kind of format examined not only lower-order visual cognitive skills but also higher-order visual cognitive skills, except for create level.

The free response format was thought of as an effective means of evaluating a variety of cognitive skills [12]. As shown in Fig. 8, however, the free response format (the sum of F1, F2, and F3) was predominantly used to examine LOCS visual tasks in four metropolises, Beijing (37.75%), Shanghai (37.93%), Hangzhou (36.67%), and Suzhou (32.88%), while HOCS visual tasks were not adequately addressed, with Beijing (10.20%), Shanghai (25.86%), Hangzhou (23.34%), and Suzhou (14.77%). Moreover, in the four metropolises' SHSEE physics, the free response format was presented mainly in the form of free verbal or numerical responses (F1), with a small number of tasks based on free graphical responses (F2) and chart completion with free verbal responses (F3). In other words, the visual representations in these items are used to simply represent the content from texts, thus contributing little to the development of students' higher-order visual cognitive skills.

V. DISCUSSION AND IMPLICATIONS

To address the issue of testing scientific visual literacy in large-scale external examinations, 12 sets of SHSEE physics of the Senior High School Entrance Examinations (SHSEE) in four Chinese metropolises were targeted in this study. As mentioned earlier, the development of higher-order visual cognitive skills is crucial for students' scientific visual literacy but the extent to which different visual cognitive skills are examined in external examinations remains unexplored. In this sense, this study has filled the gap. Specifically, we adopted the VBT to analyze 364 visual-based questions extracted from the 12 sets of the SHSEE physics from two aspects: the distribution of visual cognitive skills and the distribution of visual cognitive skills in different answer formats. The results have revealed that both LOCS and HOCS were involved in the SHSEE physics of the four metropolises. However, in these examinations, HOCS primarily concentrated on the levels of apply and analyze, with insufficient attention given to evaluate and create. Regarding the response format, it was found that multiple choice covered not only LOCS (i.e., remember, understand, apply LOCS), but also HOCS (i.e., apply HOCS, analyze, and evaluate), allowing students to demonstrate their ability of decoding visual representations in a comprehensive manner. On the other hand, the free response format placed a greater emphasis on interpreting visual representations rather than constructing them, which may be detrimental to developing students' scientific visual literacy. In view of the above common issues in the four metropolises, it is suggested that the visual-based physics questions in the SHSEE still have much room for improvement.

First of all, more emphasis should be placed on testing evaluate level and create level when designing visual-based questions. According to the MoE [5], junior high school graduates should be able to use simple diagrams or tables to describe information, use different visual representations to describe scientific inquiry results, and compare visuals to find features in them, all of which put forward high requirements for students' higher-order vi sual cognitive skills. The results showed that for the HOCS, the four metropolises focused primarily on apply and analyze, while evaluate and create were limited. For example, 7 of the 12 SHSEE physics did not examine the evaluate level or create level. According to Schönborn and Anderson [45], the evaluate level involves assessing the power, limitations, and overall quality of the visual representations, and students are required to develop the ability to decode the visual representations accordingly; the create level requires students to construct, modify, and utilize their own visual representations as part of scientific inquiry, thereby enhancing their ability to encode visual representations. If these two visual cognitive skills are set properly, the higher-order visual cognitive skills can be examined more adequately, which is conducive to the development of students' scientific visual literacy [46]. As such, it is suggested to appropriately increase the proportion of evaluate and create in future large-scale external physics examinations like SHSEE to fully assess students' understanding of a wide range of visual cognitive skills.

Moreover, more attention should be paid to testing students' ability to encode visual representations when designing visual-based questions. Based on the multimedia learning model [47], Van Meter and Garner [48] proposed that the encoding of visual representations could be divided into three steps. First, students should identify relevant key elements from the text, the selected key elements are then organized into an internal verbal representation of the text, and finally, students construct an internal visual representation of the text associated with the verbal representation and relevant prior knowledge. As the process of encoding visual representations provides students with the opportunity to foster higher-order cognitive skills such as organization and integration of material as well as metacognitive self-monitoring [49], it has received special attention in curriculum documents in recent years [5,6]. Nevertheless, the results of this study showed that free response questions were dominated by LOCS in all four metropolises. Additionally, the visual-based questions in four metropolises contained a small percentage of free graphical responses (F2) and chart completion with free verbal responses (F3). In the two relevant studies, Yeh and McTigue [8] studied standardized science tests at the elementary and middle school in the United States and found that the average proportion of F2 and F3 in visual-based questions was only 2.3% and 1%, respectively. Similarly, Unsworth and Herrington [9] examined final-year science examinations in high schools in three countries (Australia, Singapore, and New Zealand). Their study showed that the average percentage of free response questions in physics examinations that required students to construct visual representations was extremely low. In comparison, the four metropolises exhibited a slightly higher percentage of questions that required students to construct visual representations. Even so, the overall proportion of such questions in the examinations is limited. To fully exploit the advantages of free responses and reflect the curriculum's advocacy for HOCS, it is necessary to increase the proportion of free response formats in visual-based questions (particularly F2 and F3) in future examinations. Also, the proportion of HOCS in free response formats should be appropriately increased.

Finally, it is essential to develop robust frameworks for assessing the visual representations that students produce in high-stakes examinations. According to Ainsworth *et al.* [50], students need to experience continuous and diverse translations between visual and text representations to reach a deeper level of understanding. In particular, as a means to facilitate the development of students' higher-order cognitive skills, it is necessary to appropriately increase the proportion of constructing visual representations in the examination. However, our study indicated that only a small percentage of items in the SHSEE physics require students to construct visual representations. A possible explanation is that high-stakes examinations focus on marking reliability, while free graphic response (F2) questions are open-ended, making it difficult to control marking reliability. In order to assist examiners in obtaining reliable and valid marking bases for high-stakes examinations, it is necessary to develop robust frameworks for assessing the visual representations generated by students [51,52]. For example, Tang *et al.* [52] proposed an analytical framework for examining visual representations encoded by students based on sociosemiotic theory, which includes seven categories: Association, Spatial, Movement, Perspective, Modality, Connective, and Textual Contextualization. This framework has been illustrated by two examples of visual representations produced by undergraduate students in the areas of physics and chemistry [52]. Furthermore, since different analytical frameworks may emphasize different aspects, it is essential to select the appropriate framework according to the purpose (diagnostic, formative, or summative). and format (paper and pencil or online) of the examinations.

The findings of this study have implications for future studies. First, this study was conducted with 12 sets of the SHSEE physics from the four Chinese metropolises over the last 3 years. It should be very prudent to generalize the results of this study to the overall state of the SHSEE physics across the country. In the future, larger and more diverse samples could be involved to validate the results obtained in this study and to gain a more comprehensive understanding of SHSEE physics in China regarding the test of visual cognitive skills. Second, this study was intended to examine the extent to which examination items reflect the curriculum's advocacy for higher-order visual cognitive skills, but we did not specifically examine the alignment between the examination items and the intended curriculum at the level of test items. In the future, a detailed analysis of their alignment with the intended curriculum will be conducted, facilitating the identification of specific problems and issues about the science examination items. Third, this study focused on the SHSEE at the end of the compulsory education stage in China. Future researchers could focus on the College Admission Examination (usually called *Gaokao* in China) at the end of the secondary schooling. Furthermore, international research on the assessment of visual cognitive skills in external examinations is currently insufficient. To facilitate cross-country comparative studies, it is imperative to further investigate the extent to which different visual cognitive skills are tested in large-scale external examinations around the globe.

VI. ETHICAL STATEMENT

No human participant was involved in this study and all data used in this study are open to the public.

[1] M. Kim and Q. Jin, Studies on visualisation in science classrooms: A systematic literature review, Int. J. Sci. Educ. **44**, 2613 (2022).

[2] E. M. McTigue and A. C. Flowers, Science visual literacy: Learners' perceptions and knowledge of diagrams, Reading Teach. **64**, 578 (2011).

[3] Z. H. Zhang and M. C. Linn, Learning from chemical visualizations: Comparing generation and selection, Int. J. Sci. Educ. **35**, 2174 (2013).

[4] K. K. Matusiak, C. Heinbach, A. Harper, and M. Bovee, Visual literacy in practice: Use of images in students' academic work, Coll. Res. Libr. **80**, 123 (2019).

[5] Ministry of Education (MoE), *The Compulsory Education Physics Curriculum Standards (2022 Version)* (Beijing Normal University Press (in Chinese), Beijing, 2022).

[6] NGSS Lead State, *Next Generation Science Standards: For State, by State* (The National Academy Press, Washington, DC, 2013).

[7] N. D. LaDue, J. C. Libarkin, and S. R. Thomas, Visual representations on high school biology, chemistry, earth science, and physics assessments, J. Sci. Educ. Technol. **24**, 818 (2015).

[8] Y. F. Y. Yeh and E. M. McTigue, The frequency, variation, and function of graphical representations within standardized state science tests, School Sci. Math. **109**, 435 (2009).

[9] L. Unsworth and M. Herrington, Visualization type and frequency in final year high school science examinations, Res. Sci. Educ. **53**, 707 (2023).

[10] Y. Wu, The examination system in China: The case of Zhongkao mathematics, in *Selected regular lectures from the 12th International Congress on mathematical education* (Springer, New York, NY 2015), pp. 897–914.

[11] A. Crowe, C. Dirks, and M. P. Wenderoth, Biology in bloom: Implementing Bloom's taxonomy to enhance student learning in biology, CBE—Life Sci. Educ. **7**, 368 (2008).

[12] R. B. Kozma and J. Russell, Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena, J. Res. Sci. Teach. **34**, 949 (1997).

[13] N. Frederiksen, The real test bias: Influences of testing on teaching and learning, Am. Psychol. **39**, 193 (1984).

[14] M. G. Simkin and W. L. Kuechler, Multiple choice tests and student understanding: What is the connection?, Decis. Sci. J. Innov. Educ. **3**, 73 (2005).

[15] S. Ainsworth, The educational value of multiple-representations when learning complex scientific concepts, in *Visualization: Theory and Practice in Science Education* (Springer Netherlands, Dordrecht 2008), pp. 191–208.

[16] D. Guo, E. M. McTigue, S. D. Matthews, and W. Zimmer, The impact of visual displays on learning across the disciplines: A systematic review, Educ. Psychol. Rev. **32**, 627 (2020).

[17] K. F. Stanger-Hall, Multiple-choice exams: An obstacle for higher-level thinking in introductory science classes, CBE Life Sci. Educ. **11**, 294 (2012).

[18] C. D. Tippett, What recent research on diagrams suggests about learning with rather than learning from visual representations in science, Int. J. Sci. Educ. **38**, 725 (2016).

[19] J. B. Arneson and E. G. Offerdahl, Visual literacy in Bloom: Using Bloom's taxonomy to support visual learning skills, CBE Life Sci. Educ. **17**, ar7 (2018).

[20] J. Trumbo, Visual literacy and science communication, Sci. Commun. **20**, 409 (1999).

[21] B. Wei, C. Wang, and L. Tan, Visual representation of optical content in China's and Singapore's junior secondary physics textbooks, Phys. Rev. Phys. Educ. Res. **18**, 020138 (2022).

[22] Y. P. Ge, L. Unsworth, and K. H. Wang, The effects of explicit visual cues in reading biological diagrams, Int. J. Sci. Educ. **39**, 605 (2017).

[23] M. P. Cook, Graphic representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles, Sci. Educ. **90**, 1073 (2006).

[24] M. Avgerinou and J. Ericson, A review of the concept of visual literacy, Br. J. Educ. Technol. **28**, 280 (1997).

[25] B. Blummer, Some visual literacy initiatives in academic institutions: A literature review from 1999 to the present, J. Vis. Lit. **34**, 1 (2015).

[26] R. K. Lowe, Animation and learning: Selective processing of information in dynamic graphics, Learn. Instr. **13**, 157 (2003).

[27] S. E. Metros, The educator's role in preparing visually literate learners, Theory Pract. **47**, 102 (2008).

[28] K. J. Schönborn and T. R. Anderson, The importance of visual literacy in the education of biochemists, Biochem. Mol. Biol. Educ. **34**, 94 (2006).

[29] J. G. Cromley, S. M. Weisberg, T. Dai, N. S. Newcombe, C. D. Schunn, C. Massey, and F. J. Merlino, Improving middle school science learning using diagrammatic reasoning, Sci. Educ. **100**, 1184 (2016).

[30] J. Airey and C. Linder, A disciplinary discourse perspective on university science learning: Achieving fluency in a critical constellation of modes, J. Res. Sci. Teach. **46**, 27 (2009).

[31] E. G. Offerdahl, J. B. Arneson, and N. Byrne, Lighten the load: Scaffolding visual literacy in biochemistry and molecular biology, CBE Life Sci. Educ. **16**, es1 (2017).

[32] F. Al-Salmani and B. Thacker, Rubric for assessing thinking skills in free-response exam problems, Phys. Rev. Phys. Educ. Res. **17**, 010135 (2021).

[33] B. Wei and Y. Ou, A comparative analysis of junior high school science curriculum standards in Mainland China, Taiwan, Hong Kong, and Macao: Based on revised Bloom's taxonomy, Int. J. Sci. Math. Educ. **17**, 1459 (2019).

[34] B. Wei, The change in the intended Senior High School Chemistry Curriculum in China: Focus on intellectual demands, Chem. Educ. Res. Pract. **21**, 14 (2020).

[35] E. Ene and B. J. Ackerson, Assessing learning in small sized physics courses, Phys. Rev. Phys. Educ. Res. **14**, 010102 (2018).

[36] M. D. Engelhart, E. J. Furst, W. H. Hill, and D. R. Krathwohl, *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook 1: Cognitive Domain*, B. S. Bloom (David McKay, New York, NY, 1956).

[37] L. W. Anderson, D. R. Krathwohl, P. Airasian, K. Cruikshank, R. Mayer, P. Pintrich, and M. Wittrock, *A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy* (Longman, New York 2001).

[38] L. Mnguni, K. Schönborn, and T. Anderson, Assessment of visualisation skills in biochemistry students, South Afr. J. Sci. **112**, 8 (2016).

[39] P. Sinaga, I. Kaniawati, and A. Setiawan, Improving secondary school students' scientific literacy ability through the design of better science textbooks, J. Turk. Sci. Educ. **14**, 92 (2017), http://www.tused.org/index.php/tused/article/view/176.

[40] K. Krippendorff, *Content Analysis: An Introduction to Its Methodology* (Sage Publications, Inc, Thousand Oaks, CA, 2018).

[41] S. E. Stemler, Content analysis in *Emerging Trends in the Social and Behavioral Sciences: An Searchable and Linkable Resource* (John Wiley & Sons, Inc., New York, 2015), pp. 1–14.

[42] OECD, *PISA 2018 Results (Volume I): What Students Know, and Can Do* (PISA, OECD Publishing, Paris 2019), 10.1787/5f07c754-en.

[43] A. B. Cantor, Sample-size calculations for Cohen's kappa, Psychol. Methods **1**, 150 (1996).

[44] N. J. M. Blackman and J. J. Koval, Interval estimation for Cohen's kappa as a measure of agreement, Stat. Med. **19**, 723 (2000).

[45] K. J. Schönborn and T. R. Anderson, Bridging the educational research teaching practice gap: Foundations for assessing and developing biochemistry students' visual literacy, Biochem. Mol. Biol. Educ. **38**, 347 (2010).

[46] W. S. Van Rooy and E. Chan, Multimodal representations in senior biology assessments: A case study of NSW Australia, Int. J. Sci. Math. Educ. **15**, 1237 (2017).

[47] R. E. Mayer, The promise of multimedia learning: Using the same instructional design methods across different media, Learn. Instr. **13**, 125 (2003).

[48] P. Van Meter and J. Garner, The promise and practice of learner-generated drawing: Literature review and synthesis, Educ. Psychol. Rev. **17**, 285 (2005).

[49] A. Schmeck, R. E. Mayer, M. Opfermann, V. Pfeiffer, and D. Leutner, Drawing pictures during learning from scientific text: Testing the generative drawing effect and the prognostic drawing effect, Contemp. Educ. Psychol. **39**, 275 (2014).

[50] S. Ainsworth, V. Prain, and R. Tytler, Drawing to learn in science, Science **333**, 1096 (2011).

[51] H. Y. Chang, Students' representational competence with drawing technology across two domains of science, Sci. Educ. **102**, 1129 (2018).

[52] K. S. Tang, M. Won, and D. Treagust, Analytical framework for student-generated drawings, Int. J. Sci. Educ. **41**, 2296 (2019).