# Assessment of knowledge integration in student learning of galvanic cell: An interdisciplinary approach connecting physics and chemistry

Dewei Ye,[1,2] Jiali Qian,[3] Wangyi Xu,[4] Yuyang Lu,[5] Shuli Wang,[6]
Longhai Xiao,[1,†] and Lei Bao[2,*]

[1]*College of Education, Zhejiang University, Hangzhou, Zhejiang 310058, China*
[2]*Department of Physics, The Ohio State University, Columbus, Ohio 43210, USA*
[3]*School of Education, Shanghai Jiao Tong University, Shanghai 200240, China*
[4]*Jing Hengyi School of Education, Hangzhou Normal University, Hangzhou, 311121, China*
[5]*Department of Chemistry, Zhejiang University, Hangzhou, Zhejiang 310058, China*
[6]*Hangzhou Yuhang High School, Hangzhou, Zhejiang 311121, China*

(Received 16 May 2024; accepted 21 September 2024; published 15 October 2024)

The galvanic cell is a typical interdisciplinary topic, which is primarily taught in chemistry but its underpinning is closely related to physics. Student learning in galvanic cell has been extensively studied in chemistry education, which has revealed a large number of misconceptions that are difficult to change through traditional instruction. A source of the learning difficulties is that most students lack an explanatory framework to integrate the many complex phenomena and processes into a coherent knowledge system. As a result, students often rely on memorization of terms, laws, and equations in solving problems but without a meaningful understanding of their reasoning. To address the deficit in learning, this study aims to help students develop an explanatory framework for supporting an integrated knowledge structure. Specifically, a conceptual framework model for galvanic cell is developed based on the central idea from physics, which is used as the core concept for mechanistic explanations of the related chemistry concepts. Guided by the conceptual framework, an instrument is developed to assess the knowledge integration in student learning of galvanic cell. A combination of qualitative and quantitative measures was used to assess the levels of knowledge integration in learning galvanic cell with a large number of Chinese high school and college students. The assessment results show that the conceptual framework model can effectively represent the knowledge structures of students at different levels of knowledge integration. The galvanic cell conceptual framework and assessment results can provide useful resources for teachers to develop instruction for promoting knowledge integration.

DOI: 10.1103/PhysRevPhysEducRes.20.020129

## I. INTRODUCTION

As scientific research progresses, scientists increasingly integrate interdisciplinary knowledge to tackle complex research challenges [1]. This interdisciplinary approach has led to significant scientific breakthroughs [2–4] and profoundly influenced the knowledge structures within disciplines. Interdisciplinary insights have often provided essential explanatory frameworks for understanding specific fields. Nevertheless, traditional instruction still emphasizes the memorization of operational rules, laws, and equations within specific disciplines [5,6], while neglecting the closely related interdisciplinary knowledge, which may provide useful insights to develop a deeper understanding of disciplinary knowledge. This omission of interdisciplinary understanding sometimes may impede students from forming a comprehensive knowledge structure. This situation often manifests as many students being limited to solving simple problems by relying solely on the rote-based learning, leading to difficulty in working with more complex problems that require deep explanatory understanding and interconnected knowledge systems.

According to the existing research on students' problem-solving behaviors, the context dependence and connectedness of students' knowledge structure is a crucial criterion for distinguishing novices from experts [7–10]. Experts possess well-organized, hierarchically networked knowledge structures centered on central ideas [7]. These structures feature consistent, long-range links that enable abstract and deep understanding, facilitating the application of principles across different contexts [7]. In contrast, novices' knowledge structures are fragmented with limited

*Contact author: bao.15@osu.edu
†Contact author: zdlonghai@zju.edu.cn

connections, leading them to rely on memorization of problem-solving strategies associated with surface features of problems [7,11,12]. Thus, novices' ability to successful problem solving is typically restricted to familiar contexts seen in textbooks or highlighted by educators.

For many topics in science courses, the central ideas for fundamental explanations of the content often come from outside the individual disciplines, which require exploration and learning of concepts across different disciplines [13]. For example, physics often plays an important role in offering the mechanistic explanation to many phenomena in chemistry and biology. However, due to the limitations imposed by the single-discipline system in school education, science curricula often focus exclusively on discipline-specific knowledge, thereby neglecting the integration of interdisciplinary insights. Such limitations often result in students being unable to develop an integrated deep understanding of disciplinary knowledge due to the lack of explanatory central ideas [14].

Narrowing down to the teaching and learning of physics and chemistry, many topics in the two disciplines share profound relationships, where a large number of chemistry concepts are deeply rooted in physics [15–17]. Therefore, many central ideas that explain chemical phenomena are originated from physics. For example, the galvanic cell is a typical interdisciplinary topic, which is primarily taught in chemistry, but its underlying mechanism is in physics. To help students develop a deeper and integrated understanding of the galvanic cell concept, it is then helpful to introduce the related physics principles into the chemistry instruction. However, traditional chemistry instruction often neglects the central ideas derived from physics, leading to students acquiring knowledge that is only operational but lacks explanatory reasoning. Consequently, students often resort to memorizing terms, laws, and equations to solve problems, which impedes their ability to engage in deep scientific reasoning.

Recognizing the limitations of the traditional chemistry instruction in providing the fundamental explanatory mechanisms to many chemical processes and problems, this study employs the conceptual framework model of knowledge integration to introduce an interdisciplinary-based explanatory central idea in the instruction of a chemistry topic, the galvanic cell. In this knowledge integration approach, a conceptual framework of galvanic cell is developed, which determines the central idea of the concept that explains the physical chemistry phenomena and aids students in developing a coherent knowledge system. Assessment tools are then developed based on this conceptual framework to accurately gauge the student's level of knowledge integration attained, providing valuable information for instruction through assessment results. Consequently, specific conceptual pathways and connections can be emphasized during instruction to facilitate knowledge integration. Previous studies have established conceptual framework models for physics concepts and developed assessment tools based on the conceptual frameworks to assess students' knowledge integration levels [13,18–24]. In this study, we extend the existing research to develop an interdisciplinary conceptual framework using physics concepts as the central idea to support the explanatory reasoning needed for knowledge integration in learning galvanic cell in chemistry.

The conceptual framework model typically categorizes students into three developmental levels: novice, intermediate, and expertlike [13,18–24]. In learning, the learner's ideas are activated and dependent on contextual features [21,22]. Novices typically exhibit fragmented knowledge structures, often bypassing the central idea and directly matching surface features of a context with memorized terms, laws, and equations in problem solving. In contrast, the intermediate students have developed partial connections within their knowledge structures, yet their understanding of the central idea, context features, and operating rules remains incomplete. While they may navigate certain contexts more effectively than novices, they often struggle in unfamiliar or complex situations due to a lack of deep understanding of the central idea. Experts, by contrast, connect activated thoughts and related conceptual components through the central idea. They form a cohesive reasoning pathway around this central idea, thereby establishing a fully integrated knowledge structure. Therefore, experts can solve problems based on reasoning with the central idea when facing unfamiliar and complex contexts.

To develop the conceptual framework for a specific concept, it is crucial to first clearly define the central idea of the concept and its associated contextual variables and relationships [13,18–24]. The next step involves identifying conceptual pathways that link the central idea to other elements of the concept. Once established, this conceptual framework becomes instrumental in developing assessment tools designed to evaluate student's level of knowledge integration. These assessment tools generally include single-link typical questions, multilink atypical questions, and integrated-link atypical questions. Single-link typical questions are primarily composed of familiar problems, often highlighted in textbooks or during instruction, and can be resolved using memorized terms, laws, and equations. Multilink atypical questions require students to connect concepts across both familiar and unfamiliar contexts to solve problems. Integrated-link atypical questions challenge students to resolve problems by applying the central idea in unfamiliar situations. By conducting a comparative analysis of student performance on the three different types of links and context saliency question sets and integrating their performance data with findings from student interviews, it is possible to reveal students' knowledge integration levels and reasoning pathways.

To better help students understand the galvanic cell, this study aims to introduce the central idea from physics to help students construct an integrated knowledge structure for their understanding of the galvanic cell. This research is conducted with two areas of study:

Study 1: Integrate expert views and student understandings of the galvanic cell to establish a conceptual framework model of the galvanic cell and analyze students' understanding of the galvanic cell through the knowledge integration perspective by applying the conceptual framework model.

Study 2: Apply the conceptual framework model to develop an assessment instrument for evaluating the knowledge integration levels of students in learning galvanic cell.

## II. METHOD AND DESIGN

### A. Expert views and student understandings of galvanic cell

The galvanic cell is a complex interdisciplinary topic that is closely linked with chemical and physical concepts such as electrode potential, potential difference, simple electric circuits, and oxidation-reduction reactions [25,26]. From the expert's view of the knowledge, related physics concepts can provide an explanation for the driving force behind the occurrence of galvanic cell reactions. Fundamentally, the atoms composing different electrode materials exhibit differences in structures, such as the number of valence electrons and electrical layers. Consequently, their capabilities to gain or lose electrons differ, leading to uneven charge distribution when different electrode materials are immersed in a solution. This results in the establishment of electric fields and potential differences between the electrodes and the solution. In the field of chemistry, this potential difference is defined as the electrode potential.

A related concept in chemistry is the oxidation-reduction property of materials, which also reflects the capability of substances to gain or lose electrons in chemical reactions. However, the oxidation-reduction property is often presented as measured outcomes without further mechanistic explanations, leading students to merely memorize this property without meaningful understanding and reasoning. In contrast, the concept of electrode potential is constructed through the physical chemistry process based on the electron gain or loss capability of substances, providing more physically based explanations for galvanic cell reactions. Therefore, experts use electrode potential to describe the electron gain or loss capability of substances in these reactions. In addition, when electrodes are in contact, electrode materials with different electron gaining or losing capabilities also lead to different charge distributions on their contact surfaces. Consequently, the contact between electrode materials results in the generation of electric fields and potential differences across the contact surface. This potential difference is defined as the contact potential in chemistry.

Furthermore, different charge densities within solutions also give rise to electric fields and potential differences within solutions. In the field of chemistry, this potential difference is defined as the liquid junction potential. However, the contact potential and liquid junction potential are so small that they are often neglected in modeling the galvanic cell. Therefore, after assembling a galvanic cell, the difference in electrode potentials of the positive and negative electrodes gives the potential difference between the positive and negative electrodes. Under the influence of this potential difference, electric charge carriers gain directional motion, generating a current and thereby initiating oxidation-reduction reactions. This process is illustrated in Fig. 1(a). Here in physics, potential difference and electric field are two coupled concepts linked through $\Delta\phi = \vec{E} \cdot \vec{d}$ for a constant electric field $\vec{E}$ and a displacement $\vec{d}$ in space. In a circuit with a potential difference, the associated electric field will apply forces on charges, leading to directed motions of charges to form an electric current. In chemistry discussions, the explanation for creating an electric current often uses potential difference without referring to the electric field. It should be noted that in this paper, whenever a potential difference is mentioned, the associated electric field is assumed to provide the driving force for the electric current.

It is also noted that the terms "cathode" and "anode," respectively, refer to the electrodes in oxidation-reduction reactions that gain or lose electrons. Meanwhile, "positive electrode" or "negative electrode" designate the ends of a circuit with higher or lower electrical potentials. In a galvanic cell, the cathode gains electrons and has a higher potential, thus serving as the positive electrode, while the anode loses electrons and has a lower potential, thereby serving as the negative electrode. This study adopts the terminology of "positive electrode" and "negative electrode" to define the two electrodes of the galvanic cell.

The established experts' understanding of the galvanic cell demonstrates that related physics concepts can provide mechanistic explanations for terms and procedures used in chemistry. Therefore, incorporating these physics concepts into chemistry instruction on galvanic cells can potentially help students develop more meaningful reasoning, enhance their grasp of chemistry terms and procedures, and foster a deeper, integrated understanding of the concept.

To examine how instruction of galvanic cell are implemented in different countries, we analyzed high school general chemistry textbooks from China, Japan, Australia, and America that cover the topic of galvanic cells. These textbooks are designed based on the national curriculum standards of the respective countries, and therefore, the content related to galvanic cell is essentially the same

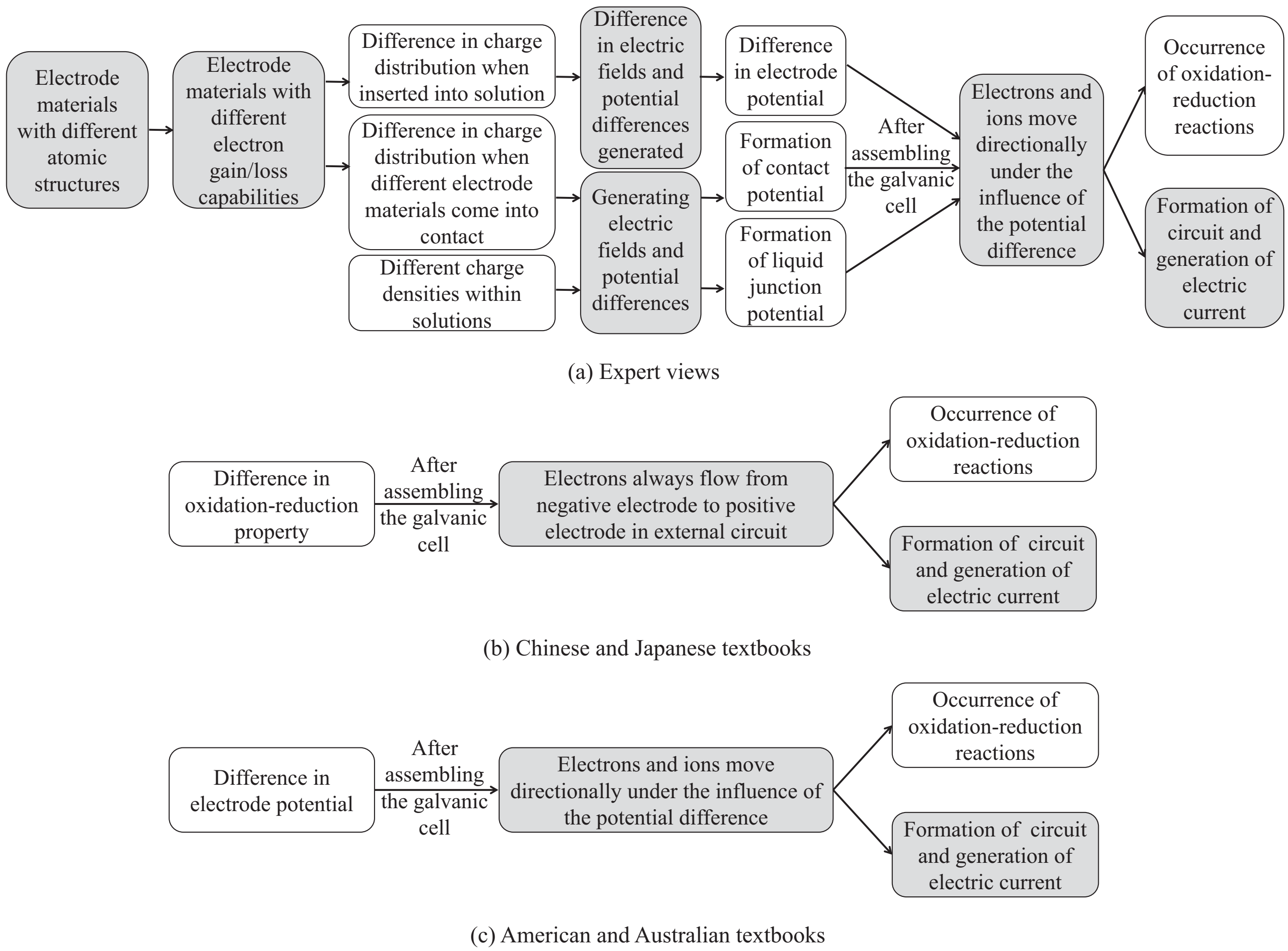


FIG. 1. Causal pathways in the teaching and learning of galvanic cell: (a) expert views on interdisciplinary-based causal pathways of galvanic cell instruction, (b) causal pathways of galvanic cell instruction in Chinese and Japanese textbooks, (c) causal pathways of galvanic cell instruction in American and Australian textbooks.

across textbooks within the same country. As a result, we selected one representative textbook from each country for our analysis.

However, upon reviewing the current instruction on the galvanic cell in chemistry textbooks, we found that the instruction often remains limited to concepts within chemistry. For example, the Chinese [27] and Japanese [28] textbooks introduce the oxidation-reduction property as the basis for determining the positive and negative electrodes without further explanation. These properties are given in a table known as the reactivity series of metals. As a result, students would rely solely on looking up and or memorizing the table of the reactivity series of metals to determine the positive and negative electrodes and chemical reactions occurring on the electrodes. Consequently, students often lack the understanding to conduct meaningful reasoning with the concept of oxidation-reduction property [29,30], which leads to students' difficulties in solving problems when the oxidation-reduction property of unfamiliar metals is presented. On the other hand, after assembling the galvanic cell, a circuit is formed. Without a proper discussion of the physics concepts of electric circuits, students are limited to memorizing the law "Electrons always flow from the negative electrode to the positive electrode in external circuit" to recall the direction of electrons and current. This instructional process is illustrated in Fig. 1(b), where the gray portion represents the physics concepts or laws within the causal pathways of galvanic cell instruction.

In the American [31] and Australian [32] chemistry textbooks, students are directly instructed to memorize a set of standard electrode potential values and use the potential relationships reflected by standard electrode potentials to determine the positive and negative electrodes, as well as the chemical reactions occurring on the electrodes. After assembling the galvanic cell, students are taught that under the influence of potential difference, charges move directionally to generate an electric current and undergo oxidation-reduction reactions, as illustrated in Fig. 1(c).

Based on our review of textbooks from four countries, it was found that after assembling a galvanic cell, chemistry textbooks in China and Japan still only require students to memorize the behaviors of electron's movement in the external circuits to solve problems. In contrast, chemistry textbooks in America and Australia introduce the concept of potential difference, which determines the direction of an electron's movement and the formation of electric current. This indicates that the instruction of galvanic cell in China and Japan remains at the level of memorization of chemistry terms, without meaningful links to the underlying physics concepts. On the other hand, the American and Australian textbooks start to link galvanic cell with physics concepts on electric circuits, which is an encouraging step toward fostering knowledge integration. Unfortunately, none of the textbooks introduce the physics concepts to fundamentally explain the underlying mechanisms of the galvanic cell reactions. Current chemistry textbooks primarily require students to memorize terms such as standard electrode potentials or the oxidation-reduction property of substances. These terms are measured outcomes that characterize the varying electron gaining or losing capabilities of electrode materials, yet they do not incorporate the underlying physics concepts in the curriculum. In addition, the textbooks require students to judge the positive and negative electrodes and the chemical reactions that occur based on standard electrode potentials or oxidation-reduction property of substances. However, there is no detailed explanation of how the different electron gaining or losing capabilities of electrode materials lead to the potential differences between different interfaces, which is the fundamental driving force behind the reactions in a galvanic cell.

Based on the analysis of the curricula on galvanic cell in current textbooks, there are apparent deficiencies in teaching the mechanisms underlying the concepts and operations of electrode potential and simple electric circuits. These deficiencies in existing curricula can lead to numerous misconceptions or learning difficulties among students toward understanding the galvanic cell reaction and galvanic cell circuits, which are reviewed and summarized next.

#### *1. Students' misconceptions about galvanic cell reaction*

The potential difference between different interfaces serves as the fundamental impetus for reactions within galvanic cell. Consequently, grasping the mechanism underlying these potential differences is a critical prerequisite for a comprehensive understanding of galvanic cell reactions. Research indicates prevalent student misconceptions regarding the mechanism of potential differences between different interfaces. For instance, some students have the misconception that positive electric charges and electrons move in opposite directions, leading to a potential difference between the two ends of a wire [33]. Furthermore, there is a tendency among students to ascribe the potential difference in a galvanic cell solely to differences in charge concentration at specific points [29] or to the type of electrolyte solution [33,34]. These misconceptions are indicative of the lack of understanding of the mechanisms of the potential differences across different interfaces.

The potential difference between the electrode material and the solution, among the different interfaces in a galvanic cell, is identified as the primary driving force for the galvanic cell reaction. Scientists define this potential difference as the electrode potential. Research reveals that students often struggle with understanding the causal mechanism of electrode potential [35]. Common misconceptions include the belief that no electric double layer is formed at the interface between the electrolyte solution and the electrode [36], and some students also failed to understand that the electrode potential cannot be measured [35]. Regarding the magnitude of electrode potential, students often believe that it is only related to the electrode material [37] while neglecting the influence of the electrolyte solution [33,38,39]. These misconceptions indicate that the instruction lacks an explanation of the mechanism leading to the potential difference between the electrode material and the solution, nor does the instruction clearly define the concept of electrode potential.

Understanding standard electrode potential is crucial for analyzing galvanic cell, as direct measurement of electrode potential is not feasible. However, a common misconception among students is to regard standard electrode potential as an absolute value [35,40,41], which is directly associated by students with the chemical properties of substances [41]. Students often believe that electrode potential can be directly measured using a voltmeter [37] disregarding the need for reference to the hydrogen standard electrode [33,40–42]. These misunderstandings indicate that instruction may not have fully explained the relationship between standard electrode potential and electrode potential. Instead, teachers may only require students to memorize the standard electrode potential values of different substances.

Determining the positive and negative electrodes based on the standard electrode potential is crucial in understanding galvanic cell reactions. In current instructional methods, however, students often rely on memorization of standard electrode potential values without understanding the underlying mechanisms. Students also tend to directly equate the standard electrode potential with the oxidation-reduction property of substances to characterize the electron gaining or losing capabilities of electrode materials in a solution. This approach leads to misconceptions among students in identifying the positive and negative electrodes and their reactants [43]. For instance, due to incorrect memorization of standard electrode potential values and oxidation-reduction property of substances, some students may mistakenly believe that the substance with the highest standard electrode potential is the reactant at the negative electrode [29,34,39,42,44]. Furthermore, students may incorrectly assume that the standard electrode potential arranges substances entirely based on their oxidizability [29,42,45].

After identifying the positive and negative electrodes and their reactants, students are expected to determine the reactions occurring in the whole cell and each half-cell. However, as revealed by research, students often incorrectly determine that after a galvanic cell reaction, the mass of the negative electrode must decrease, and the mass of the positive electrode must increase [29,34,39,42,46]. Additionally, students often believe that inert electrodes can be oxidized or reduced [29,42,47], or that no reactions occur on the surface of inert electrodes [29,34,42,47]. These misconceptions lead to incorrect determinations of the chemical reactions occurring at the negative electrode and positive electrode. Similarly, research has shown that students frequently encounter difficulties in accurately writing these reactions [42]. When writing cell or half-cell reactions, students frequently fail to consider the correct stoichiometry between positive and negative reactants or overlook the number of electrons gained or lost by the reactants [48]. The results suggest that students failed to make connections to understand the atomic structure of different substances and its effects on the reactants in cell reactions.

#### 2. Students' misconceptions about galvanic cell circuit

Due to a lack of understanding of or having a fundamental misunderstanding of the mechanism of potential differences between different interfaces, students often struggle to accurately determine the direction of flow of current and charges across different interfaces. Concerning the flow of electrons, a prevalent misconception is that electrons move from the positive to the negative electrode through the wire in a galvanic cell [39] and that electrons move from areas of high potential to areas of low potential [29,36,40]. In terms of charge distribution, students often assume that negative charges accumulate around the negative electrode, attracting cations, while positive charges gather around the positive electrode, attracting anions [33,45]. As the reaction progresses, they believe that one half-cell becomes positively charged due to an excess of cations, while the other half-cell becomes negatively charged with a comparable number of anions [34,45]. Cations and anions continue to move until they reach equal concentrations in both half-cells [29,34]. These misconceptions show that students do not understand the mechanism of the current generation in galvanic cell at the microscopic level.

On the other hand, students' limited understanding of circuit concepts and operations often hinders their ability to accurately address problems related to galvanic cell circuits. For instance, those who lack comprehension of the electromotive force concept may find it challenging to identify the positive and negative electrodes. Determining the electromotive force of a galvanic cell is crucial for identifying these electrodes and their reactants [42]. In a galvanic cell, the electromotive force is the difference between the electrode potentials of the positive and negative electrodes. A galvanic cell reaction proceeds spontaneously only if its electromotive force is greater than zero. However, a common misunderstanding among students is the belief that the electromotive force can be obtained by simply adding the electrode potentials of the positive and negative electrodes [41,42]. This misunderstanding may stem from incorrectly interpreting the formula $\varepsilon = \varphi_{+(\mathrm{Ox|Red})} + \varphi_{-(\mathrm{Red|Ox})} > 0$. Additionally, students often have misconceptions about the carriers of current in different parts of a galvanic cell, leading to errors in judging the reactants, products, and reactions. A prevalent misconception believes that positive charges are the current carriers in electrodes and wires, while electrons are the carriers in the electrolyte solution. Furthermore, many students mistakenly believe that ions do not act as current carriers in the electrolyte solution and that salt bridges provide electrons for circuit completion [29,33,34,37,38,42,45,49]. This misunderstanding may arise from teachers frequently emphasizing the concept that electric current is generated by the directional movement of electrons in conductors, leading students to believe that in all cases, electrons are the carriers of electric current. Additionally, the phrase "ions carry charges" might also be misinterpreted by students, leading them to incorrectly assume that electrons are transferred from one electrode to another through ions [45].

Based on existing literature, it is evident that traditional instruction fails to adequately emphasize the interdisciplinary central idea and related concepts from physics, leading to numerous challenges for students in understanding galvanic cell. Therefore, it is crucial to establish a conceptual framework model of galvanic cell built upon the interdisciplinary central idea from physics. This explanatory conceptual framework will assist students in developing a comprehensive and coherent understanding of galvanic cell.

### B. The conceptual framework model of galvanic cell

#### 1. Develop a conceptual framework model of galvanic cell

The most critical step in constructing this framework involves identifying the central idea of a topic. This central idea, derived from expert views, forms the cornerstone of the concept, elucidating its fundamental and causal relations with other knowledge components. Subsequently, relevant contextual variables, relations, and reasoning processes are mapped in connection to the central idea.

This process creates a network of nodes and connections within the conceptual framework.

The central idea in the conceptual framework gives the most fundamental concept of this specific topic. Previous studies, including research on misconceptions, constructed the knowledge structure of galvanic cell with elements primarily within the chemistry domain of knowledge, which limited the capacity for offering a deep conceptual explanation of the physical mechanisms underlying the chemical reactions of a galvanic cell. This led to students' understanding of many concepts related to galvanic cell being composed of largely operational rules rather than conceptual explanation. To address this issue, we assembled an interdisciplinary team of experts consisting of one physics professor, two chemistry professors, one Ph.D. student of physics, and one Ph.D. student of chemistry. The team analyzed the high school textbooks from various countries that cover the topic of galvanic cell and incorporated the explanations from university physical chemistry textbooks. This interdisciplinary approach allowed us to use appropriate physics concepts to explain the chemical reactions and operational rules that are typically taught in chemistry in learning the galvanic cell. The expert team members conducted content analysis and validation with different textbooks, thereby developing the central idea and constructing the conceptual framework presented in this paper.

Based on the expert views of the content knowledge, the central idea of the galvanic cell was established as follows:

> The differing atomic structures of electrode materials result in differences in their abilities in gaining or losing outer-shell electrons. In addition, a key property of solutions is that they have different charge densities within them. These properties will lead to differences in charge distribution when different materials are in contact with each other, and uneven charge distributions between electrodes, between electrodes and solutions, and within solutions will generate electric fields (and potential differences). The electric fields apply forces on charge carriers to form directional motion that leads to electric current.

Building off the central idea and the operating rules and reasoning processes of experts and novices, a conceptual framework of the galvanic cell is developed and shown in Fig. 2.

As depicted in Fig. 2, the central idea is at the top of the conceptual framework, signifying the foundational core understanding of the knowledge structure. The bottom layer of the framework encompasses contextual features and variables, delineating specific aspects of the chemistry context and crucial elements of galvanic cell problems. This layer primarily comprises elements such as electrode material, type of electrolyte solution, and concentration of the electrolyte solution. The intermediate layer consists of intermediate reasoning processes and operational rules. These are either directly connected to the central idea within the expert knowledge structure or associated with misconceptions commonly observed in novices, as influenced by the context. Task goals are the questions that students need to answer. Since there are many different types of task goals that questions may be constructed with, only a few typical ones are listed in Fig. 2, which include the task goals related to electrode reactions, the task goals related to determining potential relationships in the galvanic cell circuit abbreviated as "$\varphi$"; and the task goals for determining the direction of current in the galvanic cell abbreviated as "$I$." Other task goals are indicated by dots. By combining these layers and task goals, the conceptual framework establishes the structure for analyzing how learners solve galvanic cell problems.

The arrows connecting different contexts, concepts, and goals are used to represent possible reasoning pathways for learners which are derived from prior research on expert-novice differences [13,18–24]. The solid arrows illustrate the reasoning pathways that experts utilize in problem solving, which can link all knowledge components to the central idea. In contrast, dashed arrows show the reasoning pathways of novices, which mostly connect surface features with misconceptions and memorized operations and solutions. For example, novices may memorize the claim that "electrons always flow from negative electrode to positive electrode in external circuit," but some students may mistakenly remember it as "electrons always flow from the positive electrode to the negative electrode in external circuit" [39], leading to failures in solving related problems. In addition, some students may also hold the misconception that the mass of the negative electrode must decrease, and the mass of the positive electrode must increase [29,34,39,42,46], leading to incorrect determination of the reactants at the positive and negative electrodes. In contrast, intermediate students often exhibit a combination of some part of expert and novice pathways, indicating that they have not yet fully integrated their knowledge to apply the central idea consistently. The empirical validations of these pathways and the entire conceptual framework are performed through quantitative assessments and interviews, which are discussed in later sections of this paper.

To gain a deep understanding of the galvanic cell, students must grasp the causal mechanism of the potential difference between interfaces of different materials in a galvanic cell. The central idea at the top of the conceptual framework elucidates this mechanism. The potential difference in a galvanic cell includes three main components: the potential difference between the electrode and the electrolyte solution, the potential difference between two electrodes that are in contact, and the potential difference across the interface between solutions

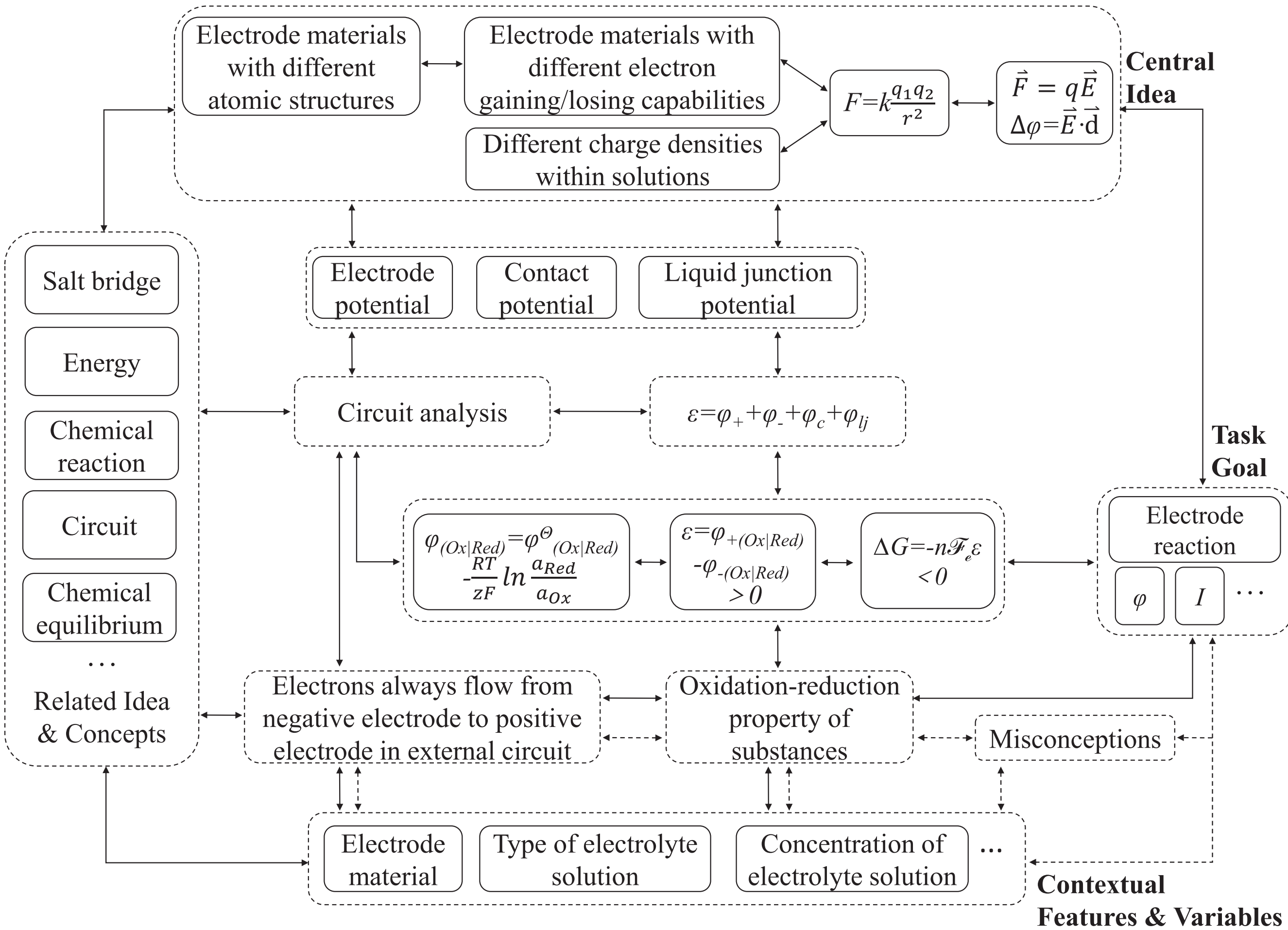


FIG. 2. The conceptual framework of galvanic cell. Solid arrows indicate the experts' conceptual pathways, and dashed arrows indicate the novices' pathways.

of different solutes or two solutions with the same at different concentrations. According to Nernst, differences in the atomic structures of substances lead to two opposing tendencies when a metal rod is immersed in its salt solution. First, metal ions from the metal atom dissolve into the salt solution due to thermal motion and the influence of polarized water molecules, a process that is more pronounced in metals with a stronger ability to lose electrons. Second, metal ions in the solution deposit onto the metal surface due to attractive forces by free electrons in the metal, a tendency that is more significant in metals with a weaker capability of losing electrons. When the processes of dissolving and depositing reach a balance, an electric double layer is formed, and the resulting potential difference is known as the electrode potential difference. Due to the different atomic structures of the electrode materials, electrons must overcome different amounts of energy as the surface barriers when escaping from the surfaces of different substances, resulting in uneven distribution of electrons, which generates a potential difference across the contact surface between two electrodes, known as the contact potential. Meanwhile, across the interface formed by two solutions of different solutes or with different concentrations of the same solute, a potential difference arises due to the different charge densities within solutions. This potential difference is known as the liquid junction potential.

The intermediate layer of the conceptual framework represents the reasoning process, which is intricately linked to the central idea within the expert knowledge structure. Researchers defined three potential differences in the galvanic cell as electrode potential ($\varphi_+$ and $\varphi_-$), contact potential ($\varphi_c$), and liquid junction potential ($\varphi_{lj}$). A thorough understanding of the mechanisms of these potential differences enables students to analyze the states and changes of electric potential at different points in the galvanic cell circuit. This analysis is crucial for determining the direction of current flow, electron movement, and ion movement. Additionally, these potential differences collectively contribute to the electromotive force of a galvanic cell. It is important to note that the contributions of contact and liquid junction potential to the overall electromotive force are often small and thus are often ignored. According to the equation $\Delta G = -n\mathfrak{F}e\varepsilon < 0$, spontaneous reactions in a galvanic cell occur only when the electromotive force is greater than zero. Therefore, the potential difference between the positive and negative electrodes of a galvanic cell can be identified by subtracting the negative electrode potential from the positive electrode potential. However, individual electrode potentials cannot be directly measured or theoretically calculated. Instead, the electrode potential of the hydrogen electrode in the standard state is defined as 0, and different standard electrode potentials are arranged in a sequential table to reflect the order of reactions. Consequently, the magnitude of the standard electrode potential is often provided in tasks to enable students to determine positive and negative electrodes, and reactions, or describe changes in electrode potentials using the Nernst equation. Furthermore, by elucidating the causal mechanism of the potential difference between different interfaces, students can analyze the circuit by reasoning through the potential difference across different points of a galvanic cell. Additionally, in many textbooks, instruction often simplifies circuit analysis to the law that "Electrons always flow from negative electrode to positive electrode in external circuit" and equates the numerical patterns of standard electrode potentials with the oxidation-reduction property of substances.

Conversely, as depicted by the dashed arrows, novices exhibit weaker connections between different nodes in the conceptual framework, resulting in a fragmented knowledge structure. This knowledge structure usually arises from their misconceptions, rote memorization, and superficial understanding of matching context variables with terms, laws, and equations. As shown in the conceptual framework of galvanic cell, novices frequently depend on familiar electrode materials, electrolyte solutions, and other given contexts to tackle problems. Moreover, they tend to memorize the reactivity series of metals to solve problems, rather than fully understanding the central idea.

It is worth noting that in the existing literature, a popular method for analyzing conceptual networks is through concept maps [50], which are visual representations showing the relationships and links among key elements (definitions and terms) related to a concept, as well as the connections between different concepts. At first glance, a concept map might seem similar to a conceptual framework, but there is a key distinction [24]. Concept maps do not illustrate the involvement of contextual configurations and conditions or their connections to various correct and incorrect conceptual ideas and operational tools, which are crucial elements forming the various conceptual pathways a learner can develop during learning. In contrast, the conceptual framework emphasizes these contextually manifested conceptual pathways, which are used to model the different aspects of students' developmental stages in knowledge integration.

### *2. Using the conceptual framework to model student understanding*

Based on the conceptual framework, further modeling can be done on the reasoning pathways behind students' misconceptions documented in the literature, thereby determining students' level of knowledge integration. Previous studies [18–24] have identified students' knowledge states through large-scale tests and interviews. These studies have found that students at different levels of knowledge integration exhibit differing performances and reasoning pathways in problem solving. In this study, the conceptual framework of galvanic cell is applied to analyze students' learning behaviors, which reveal three distinctive levels of knowledge integration:

Novice level: Students at this level typically exhibit fragmented knowledge structures, where concepts are only partially connected. They often rely on memorizing examples and formulas, which are tied to specific familiar context variables, and are not able to engage in meaningful reasoning. This tendency is particularly evident in solving galvanic cell problems. Novices often resort to memorization of the reactivity series of metals or rely on their existing misconceptions to solve problems. For example, when zinc and copper are used as electrodes with copper nitrate ($Cu(NO_3)_2$) electrolyte solution, novices can easily identify the positive electrode reactant. However, novices may find themselves struggling to solve problems when presented with unfamiliar contexts, such as when electrodes labeled X and Y are used in $Y^{2+}$ electrolyte solution, with X being more easily reducible than Y. In such cases, novices often struggle to correctly determine the positive electrode's reactant, although the concept behind this decontextualized example is identical to the ones that students can successfully solve when presented in a familiar context (i.e., using $Cu(NO_3)_2$ instead of $Y^{2+}$ electrolyte solution). The results suggest that students were memorizing the problem-solving outcomes that are tied to the surface features and lacked the basic conceptual understanding underlying the problem.

Intermediate level: Students at this level exhibit a deeper understanding of concepts than novices though they still rely on memorized examples and formulas for problem solving. Their knowledge structures display more sophisticated connections, akin to those observed in experts. Most students at this level are capable of solving problems based on the oxidation-reduction property of substances. Some intermediate students comprehend that a galvanic cell reaction can occur only when the electromotive force exceeds 0. Nonetheless, they often incorrectly add up the electrode potentials of the positive and negative electrodes based on the formula $\varepsilon = \varphi_{+(\mathrm{Ox|Red})} + \varphi_{-(\mathrm{Red|Ox})} > 0$, to assess reaction feasibility [41,42]. Additionally, some intermediate students are able to analyze electric potential differences in external circuits using their physics knowledge. They can determine the direction of current flow between electrodes and employ equations such as $\varepsilon = \varphi_{+(\mathrm{Ox|Red})} - \varphi_{-(\mathrm{Ox|Red})} > 0$ for problem solving. However, they often fail to understand the causal mechanisms of the three types of potential differences in galvanic cell. For example, they might recognize only the potential difference across the external circuit, thereby overlooking potential differences between electrodes and solutions or between different solutions. This indicates a partial understanding of the central idea.

Expertlike level: Students at this level exhibit a comprehensive understanding of the central idea, which is well connected to the intermediate reasoning processes, related concepts, and contextual variables to form an integrated knowledge structure. In the context of the galvanic cell, these students proficiently grasp the causal mechanism of the potential difference between different interfaces. They can explain the internal processes of galvanic cell reactions and fully comprehend the potential differences across different interfaces. For example, students at this level can determine the charge distribution between a liquid-metal interface and, based on this distribution, assess the potential differences throughout the galvanic cell circuit. They can also infer the direction of electron and current flow, as well as the galvanic reactions, based on the potential differences across different component interfaces throughout the galvanic cell circuit. This profound understanding empowers them to proficiently solve unfamiliar and complex problems that require the use of the central idea of galvanic cell.

These three levels represent the typical progression of students' knowledge integration. In this study, the design of assessment questions will be guided by the conceptual framework and the delineation of different levels of student understanding, thereby evaluating the knowledge structure of students regarding galvanic cell. This study will employ a combination of quantitative and qualitative methods. The quantitative data will be used to develop a quantitative scale for identifying students' levels of knowledge integration. The qualitative data will be used to identify further details of students' reasoning pathways to provide confirmative validation of the knowledge integration levels.

### C. Assessment design

Several assessments have been developed in the literature to evaluate students' conceptual understanding of galvanic cell [38,51–54]. However, these assessments primarily aim to identify students' misconceptions rather than to assess their knowledge integration. The assessment developed in this study expands its objective to include the evaluation of students' knowledge integration levels. Previous studies [18–24] have demonstrated the effectiveness of assessing students' knowledge integration based on the conceptual framework model. These assessment items differ in contextual saliency and knowledge connectedness which provide useful features to identify different levels of knowledge integration.

Previous studies have shown that questions designed with typical and atypical contexts provide an effective strategy for assessing knowledge integration [18–22]. Typical questions are those that students frequently encounter in class, textbooks, and homework, and they can usually be solved by memorizing terms, laws, and equations. In contrast, atypical questions are those that students are unfamiliar with and require the use of the central idea.

In addition, the connectedness within students' knowledge structures is also an important aspect to determine levels of knowledge integration. To measure the connectedness, the knowledge integration rubric developed by Linn *et al.* [55–58] is incorporated into the design of the assessment questions used in this study. The rubric of Linn *et al.* categorizes students' knowledge structure into six levels: ranging from low to high: no answer, no link, partial link, full link, complex link, and systemic link [57]. "No answer" indicates that the student lacks understanding of the concept entirely; "No link" suggests that the student either does not understand or misunderstands a certain concept; "Partial link" signifies that the student understands a concept but cannot link it to another concept; "Full link" indicates that the student comprehends a concept and can connect it with another related concept; "Complex link" implies that the student can link a concept with several related concepts but still has some concepts that cannot be linked; "Systemic link" demonstrates that students can link a concept with all related concepts to form a systematic understanding.

Based on the existing knowledge integration assessment methods, the Galvanic Cell Knowledge Integration Assessment (GCKIA) is developed, which includes 27 questions in three question sets with different link types and contextual saliency [19,20]. In this study, three link types are used in the question design, which include single link, multilink, and integrated link. The single-link category retains the elements of "partial link" from the original rubric, signifying a knowledge structure with minimal connections to the central idea. The multilink category combines "full link" and "complex link" elements, reflecting some isolated associations to parts of the central idea. Finally, the integrated link corresponds to the "systemic link," where knowledge components are firmly tied to the central idea. The types of "no answer" and "no link" were not present in the high school and university student populations and were therefore excluded.

When combined with the two contextual categories (typical and atypical), the design can extend into six total combinations of link type and contextual saliency. The primary goal of the assessment is to investigate students' grasp of the central idea, for which questions in atypical contexts with integrated links would provide the most insightful differentiation. Additionally, within the current

TABLE I. Questions included in different question sets and expected performance of students at different levels of knowledge integration.

| | Single-link typical Q2, Q5, Q6, Q10, Q12, Q16, Q18, Q21, Q24, Q25 | Multilink atypical Q3, Q8, Q11, Q14, Q17, Q20, Q23, Q27 | Integrated-link atypical Q1, Q4, Q7, Q9, Q13, Q15, Q19, Q22, Q26 |
|---|---|---|---|
| Novice | Sometimes successful | Rarely successful | Rarely successful |
| Intermediate | Mostly successful | Sometimes successful | Rarely successful |
| Expertlike | Always successful | Always successful | Mostly successful |

curriculum on galvanic cells, multilink and integrated-link questions are typically unfamiliar to students, whereas single-link questions are generally in familiar contexts. Consequently, it is determined that only three categories of the six combinations, which include single-link typical, multilink atypical, and integrated-link atypical, are aligned with the instructional scope and are adequate for the objective of the assessment. Similar designs have been used in previous studies and have been shown to be effective in measuring students' understanding of the central idea, a key indicator for determining knowledge integration levels [19,20]. Based on the designs, three corresponding sets of questions were developed and shown in Table I. The predicted academic performances of students with different knowledge integration levels in completing these question sets are also presented in Table I.

The single-link typical question set consists of ten questions. In this type of question, students are required to establish a single connection between specific contextual features and related terms, laws, and equations, which is shown with an example in Fig. 3(a). To solve this problem, students only need to recall the oxidation-reduction property among four metals or compare the metal activity between a certain metal and hydrogen and then draw a direct conclusion. Based on the literature, novice students may have occasional success in solving these single-link questions when the questions resemble memorized examples, however, they still have a large tendency to use their misconceptions, leading to incorrect answers. Conversely, expertlike and intermediate-level students are typically capable of successfully answering these questions.

The multilink atypical question set contains eight questions. Solving these questions requires students to establish connections between contextual features and multiple knowledge components that are partially connected to the central idea. Unlike the single-link typical questions, multilink atypical questions require the activation and correct application of multiple local links such as the standard electrode potential, circuit analysis methods, and electromotive force, to solve problems. An example of this type of question is shown in Fig. 3(b), where students not only need to utilize the concept of oxidation-reduction property to determine the positive and negative electrodes but also need to compare the potential differences between two points in the external circuit of the galvanic cell to determine the positive and negative electrodes. Although these questions are more intricate than single-link typical questions, they can still be resolved using memorized problem-solving techniques without the need to apply the central idea. Since novices typically rely on memorizing the reactivity series of metals, they often find it challenging to grasp concepts like standard electrode potential and potential difference, which limits their ability to solve such problems.

The integrated-link atypical question set contains nine questions. This type of question requires an integrated understanding based on the central idea. The ability to successfully apply the central idea to unfamiliar situations serves as a crucial indicator of whether students have attained an integrated knowledge structure. An example of this type of question is shown in Fig. 3(c). To solve this problem, students must analyze the charge distribution of

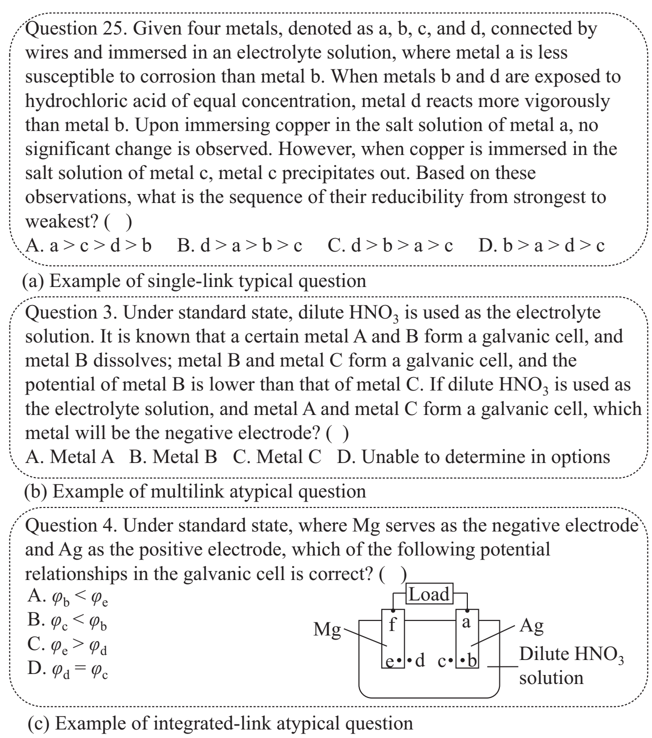


FIG. 3. (a) Example of single-link typical question, (b) example of multilink atypical question, (c) example of integrated-link atypical question.

three interfaces based on the central idea to determine the potential difference across each interface. Novice students often resort to guessing when working with such questions due to a lack of understanding of the central idea. Intermediate students also encounter difficulties in solving these problems since they have not yet developed a complete understanding of the central idea with an integrated knowledge system. Conversely, expertlike students, drawing upon their understanding of the central idea, can proficiently navigate through integrated-link atypical questions.

According to the expected performance of students on different question sets in Table I, novices mostly rely on memorized procedures of applying related terms, laws, and equations to solve problems. Therefore, they can only solve a limited number of single-link typical questions in familiar contexts and often fail to solve multilink atypical questions and integrated-link atypical questions. Intermediate-level students, who have developed more connections among concept elements, related terms, laws, and equations, can solve most single-link typical questions, and some multilink atypical questions. However, these students still lack a good understanding of the central idea and will usually fail on integrated-link atypical questions. Expertlike students are able to develop an integrated knowledge structure around the central idea, allowing them to successfully solve single-link typical questions, multilink atypical questions, and most integrated-link atypical questions. By comparing students' performances on the different question sets in GCKIA, we can then categorize students into different knowledge integration levels. Furthermore, the reasoning pathways exhibited by students during interviews provide qualitative evidence to validate whether the quantitative determination of students' knowledge integration levels aligns with the expected behaviors.

The validity of the GCKIA is evaluated in two areas including content validity and face validity. Throughout the development process, a team of experts in physics and chemistry, consisting of three professors and two graduate students, critically evaluated and refined the assessment questions to establish content validity. The design went through cycles of development and revision in multiple iterations of piloting and feedback by additional faculty and graduate students in the research institutions of the authors. The content of the final version of the assessment has been agreed to be scientifically accurate and valid for measurement by the team of designers and evaluators.

For establishing face validity, 11 high school students and 10 college students were interviewed during the piloting-revision phase of the development. In the interviews, students were asked to first explain their understanding or interpretation of the questions and the related content and then give their answers and reasoning. The analysis of the interviews primarily targeted if the students were able to accurately understand the questions and the intentions of the assessment rather than their correct or incorrect conceptions of the content. Two researchers evaluated the students' comprehension, coding their understanding into a dichotomous yes or no. The common consistency between the two researchers was 0.93 with a Cohen's Kappa at 0.83, indicating that the independent coding results of the two researchers had good interrater reliability [59]. For any coding discrepancies, the two raters discussed with a faculty member to reach a final consensus on the accuracy of students' understanding of the questions. The results showed that the correct understanding rate of the questions by the interviewed students was 0.89, indicating a solid agreement between students' understanding of the questions and the intended assessment designs [59].

### D. Data collection

Using the GCKIA, data were collected from both high school and college students. The high school students were in 11th grade from five classes in three medium-ranking high schools located in the southwest and southeastern provinces of China. College students are third-year chemistry majors from five universities in China, all of which have a national ranking among the top 150 in China. Prior to participating in the assessment, high school students had learned the basic knowledge related to galvanic cell, including electrodes, oxidation-reduction reactions, and simple circuits through their high school chemistry courses in the previous semester. The college student samples come from majors including chemistry, materials chemistry, and chemical engineering. Before taking part in the assessment, college students had learned advanced content knowledge about galvanic cell through their physical chemistry courses in the previous semester. The college-level content includes the causal mechanism of potential differences between different interfaces, electrode potential, and the Nernst equation. The time period between learning the content knowledge and the assessment is approximately 5 months for both the high school and college student populations.

The data collection was approved by the local ethics committees of the involved institutions based on the ethical policies established in China. Through the process of informed consent, all the participants were aware of the aim and content of the assessment and interview, as well as the procedures. The assessment and interview moderator informed the participants that there were no risks involved in their academic performance and their anonymity was ensured. The participants were allowed to act autonomously and to express their right to self-determination.

The primary aim of this study is to assess the extent to which high school students, who have completed a basic course on galvanic cell, and college students, who have undergone further study of galvanic cell content in physical chemistry courses, may have developed a comprehensive understanding of galvanic cell. Statistical significances of comparisons between student groups (high school and college) and question sets are determined with two-way analysis of variance (ANOVA) and further explored through one-way ANOVA and $t$ tests. Complementing the quantitative research, 25 in total, combining both high school and college groups, were randomly selected for "think-out-loud" interviews. The purpose of these interviews is to gain deeper insights into students' problem-solving approaches and reasoning pathways and to provide confirmative evidence for validating the knowledge integration levels determined with the quantitative assessment.

The interviews with high school students took place the day after the assessment, while interviews with college students were conducted within 1 week after the assessment. Each interview session lasted approximately 30 min, and all interviews were recorded and transcribed for subsequent analysis.

Valid data were extracted from the students involved in the assessment, yielding 492 data points, with 247 from high school students and 245 from college students. These data were utilized to assess different properties of the assessment, including reliability, difficulty, discrimination, and point-biserial correlation (Pb-r), for establishing the basic validity and reliability of the instrument. The results showed that Cronbach's $\alpha$ is 0.738, indicating that the assessment has good reliability ($> 0.65$) [60]. Furthermore, the overall difficulty and discrimination values are 0.573 and 0.421, respectively, falling within the satisfactory ranges [61]. The point-biserial correlations demonstrated a good alignment between individual items and the overall assessment. Further details of these statistical results are provided in Table S1 in the Supplemental Material [62].

In addition to classical statistics, Rasch analysis has been widely utilized to evaluate the effectiveness of assessment. This study used the Winsteps software package to conduct a unidimensionality Rasch analysis of the assessment [63,64]. When applying the Rasch model for measurement, assessment tool quality is typically verified through three aspects: reliability, validity, and difficulty. Further details of the Rasch results can be found in Supplemental Material Figure S1, Table S2, and Table S3 [62].

The validity index primarily includes unidimensionality and item fit (infit or outfit). A scree plot of the principal component analysis (PCA) eigenvalues is provided in Supplemental Material [62], indicating a relatively strong primary principal factor. This suggests that the evaluation data meet the criteria for unidimensionality and can be analyzed using the Rasch model. Further analysis using the PCA method of the residuals of the Rasch model revealed a residual eigenvalue of the first factor at 1.9 (falling within the reasonable range of 1.4–2.1), indicating that the scale can be considered one dimensional. Utilizing the one-dimensional Rasch model for analysis, the research findings indicated that the internal fitting index (infit MNSQ) and the external fitting index (outfit MNSQ) are at 1.00 and 0.99, respectively (the acceptable fitting index falls between 0.5 and 1.5, where 1.00 indicates the best fit) [65–67], signifying a good fit has been achieved.

The reliability index mainly includes item and person reliability and separation. The optimal value for both person and item reliability is 1.0 [66,67]. The results indicate person reliability of 0.72 and item reliability of 0.99, demonstrating good reliability for measures of both person ability and item difficulty (the value range is between 0 and 1, with higher values indicating better reliability) [66,67]. The root mean square error (RMSE) is used to estimate the error of person ability and item difficulty. The person and item RMSE are 0.47 and 0.11, respectively, indicating small estimation errors for both person ability and item difficulty (the value range is between 0 and 1, with lower values indicating better accuracy). The separation index provides a measure of discrimination and is related to reliability. An item separation value greater than 3 (item reliability $>0.9$) indicates good differentiation of item hierarchy [68,69]. In this assessment, the item separation is 8.53 (item reliability is 0.99), indicating strong discrimination of item difficulty and high reliability. The value of person separation is greater than 2 (personreliability $> 0.8$), indicating that the subjects' abilities are well distinguished. For this assessment, the person separation is 1.62 (person reliability is 0.72), which is close to 2, suggesting acceptable differentiation between subjects' abilities [68,69]. The small value of person separation may be attributed to the majority of students being at the intermediate level, with only a few students at the novice and expertlike levels, thus not fully demonstrating the range of students' abilities.

In the Rasch model, the average difficulty is set to 0. The research results indicate that the person's ability value is 0.39. The similarity between the difficulty value and ability value suggests that students' abilities and item difficulties are appropriately matched. According to Rasch analysis, it can be concluded that the assessment features are generally satisfactory and meet expectations.

## III. ANALYSIS AND RESULTS

### A. The quantitative research results

#### *1. Assessment performance of different groups of students*

The performance of high school students and college students on different question sets is shown in Table II. Two-way ANOVA was utilized to compare performance differences between student groups (high school and college) and question sets. Significant main effects were observed for both student groups $[F(1, 1470) = 298.408, p < 0.001, \eta^2 p = 0.169]$ and question sets $[F(2, 1470) = 492.108, p < 0.001, \eta^2 p = 0.401]$. Additionally, a significant interaction effect was identified between student

TABLE II. Summary of high school and college students' performance in different question sets. The mean scores and standard errors are in percentage scale.

| | High school ($N = 247$) | | College ($N = 245$) | |
|---|---|---|---|---|
| | MEAN | SE | MEAN | SE |
| Single-link typical | 69.55 | 1.10 | 81.47 | 0.98 |
| Multilink atypical | 44.33 | 1.24 | 67.55 | 1.29 |
| Integrated-link atypical | 30.81 | 1.10 | 45.85 | 1.35 |
| Total | 49.17 | 0.87 | 65.47 | 0.96 |

groups and question sets [$F(2,1470) = 12.122, p < 0.001, \eta^2 p = 0.016$].

Further analysis using one-way ANOVA was conducted to examine performance differences among high school students on three different question sets and differences among college students on three different question sets. The results revealed significant differences among high school students on three question sets [$F(2,738) = 289.964$, $p < 0.001$], with pairwise $t$ tests indicating clear distinctions between each question set [$t_{(\mathrm{S,M})}(247) = 18.429$, $p < 0.001, d = 1.36$; $t_{(\mathrm{M,I})}(247) = 9.652, p < 0.001$, $d = 0.73$; $t_{(\mathrm{S,I})}(247) = 33.731, p < 0.001, d = 2.23$]. Similarly, significant performance differences were found among college students on three different question sets [$F(2,732) = 217.815, p < 0.001$], with pairwise $t$ tests revealing notable differences between each question set [$t_{(\mathrm{S,M})}(245) = 13.932, p < 0.001, d = 0.78$; $t_{(\mathrm{M,I})}(245) = 15.283, p < 0.001, d = 1.05$; $t_{(\mathrm{S,I})}(245) = 26.955, p < 0.001$, $d = 1.94$]. These results indicate that the design incorporating link type and contextual saliency effectively distinguishes students at different performance levels.

The performances of high school students and college students on different question sets were further compared. The $t$ tests revealed a significant difference in total scores between college students and high school students [$t_{(\mathrm{C,H})}(490) = 12.558, p < 0.001, d = 1.13$]. Additionally, significant differences were observed across the three question sets: the difference between college students and high school students on single-link typical questions [$t_{(\mathrm{C,H})}(490) = 8.078, p < 0.001, d = 0.73$], the difference on multilink atypical questions [$t_{(\mathrm{C,H})}(490) = 12.918$, $p < 0.001, d = 1.16$], and the difference on integrated-link atypical questions [$t_{(\mathrm{C,H})}(490) = 8.640, p < 0.001$, $d = 0.78$]. These results indicate that college students outperformed high school students in both total scores and across different question sets. The result suggests that college students have gained a deeper understanding of galvanic cell content through their studies in physical chemistry courses. However, the poorer performance of college students on integrated-link atypical questions suggests that they have yet to achieve a fully integrated knowledge structure.

### *2. Knowledge integration levels based on performances across different question sets*

Based on previous research [13,18–24], a key factor in determining knowledge integration is students' understanding of the central idea. In the design of GCKIA, the integrated-link atypical questions are specifically targeted to probe students' understanding of the central idea of galvanic cell, which provides an assessment marker for determining knowledge integration levels. In assessment applications, it is often useful to project a developmental scale of students' knowledge integration based on students' total scores. For this purpose, trends of students' performances on the three question sets are plotted across different total scores as shown in Fig. 4. The histogram illustrates the frequency distribution of students across different total scores, while the trend lines depict the average scores on the three question sets.

The frequency of total score distribution is shown as a bar chart in the background, with the absolute count of students falling into each range shown on the right axis.

As a tradition defined in the Chinese examination passing standards, a score of 60.00% serves as the passing criterion to determine whether students have achieved basic proficiency in solving a particular question set. As depicted in Fig. 4, students with total scores below 40.00% exhibit lower average scores across all question sets. They are only able to correctly solve a few single-link typical questions while failing to correctly answer most multilink and integrated-link atypical questions. This suggests that these students lack a fundamental understanding of the central idea. They primarily rely on the surface features of the context and memorized rules such as the reactivity series of metals to solve problems, leading to poor performances on questions designed with more complex relations and in unfamiliar contexts. Therefore, these students are classified as novices in terms of knowledge integration.

Students with total scores ranging from 40.01% to 80.00% can be further classified into two groups. Those scoring between 40.01% and 60.00% could solve most of the single-link typical questions by understanding the concept of oxidation-reduction but struggled with many multilink and most integrated-link atypical questions. On the other hand, students scoring between 60.01% and 80.00% were able to successfully solve most multilink atypical questions by connecting contexts with concepts such as standard electrode potential, potential difference, and electrode potential. However, their ability to solve integrated-link atypical questions is still weak, indicating a lack of a solid understanding of the central idea. Based on students' performances on the different question sets and their implied understanding of the central idea, we can generally categorize the students with total scores between

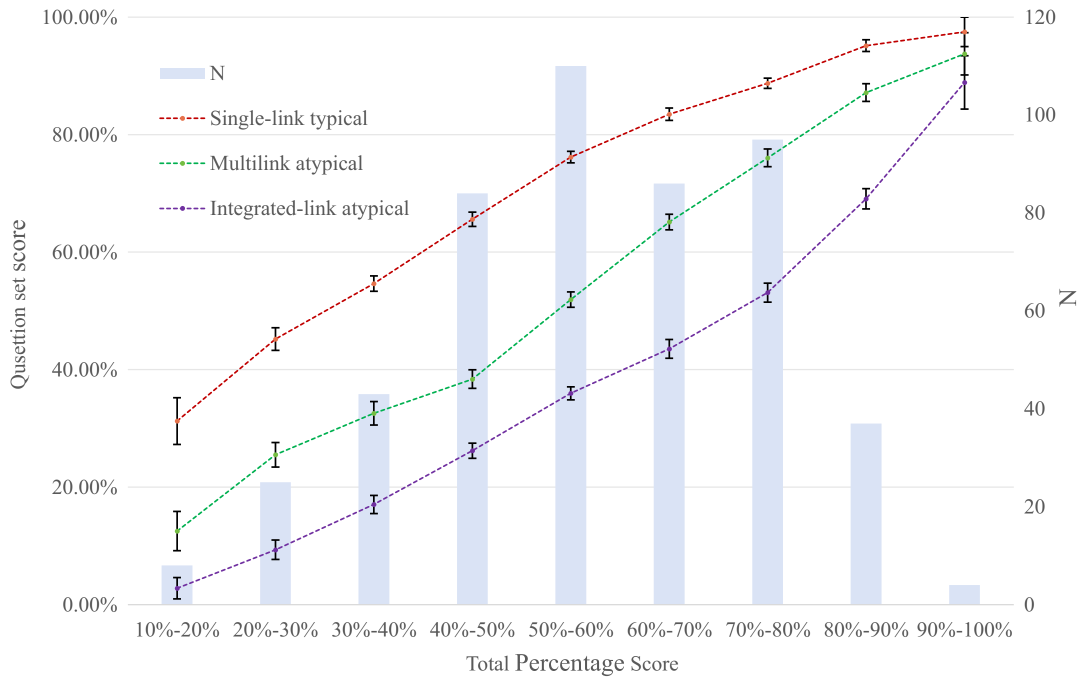


FIG. 4. Comparisons of students' performances on the three question sets. The lines show the trends of mean scores of the different question sets for students at different total score ranges with error bars denoting the standard errors. The background bar chart displays the frequency distribution of students at different total score ranges. The vertical axis on the left indicates scores on question sets for students in each total score range, while the vertical axis on the right shows the absolute count of students in each total core range.

40.01% and 80.00% at the intermediate level of knowledge integration, which can be further classified into a lower intermediate level for total scores between 40.01% and 60.00% and an upper intermediate level for total scores between 60.01% and 80.00%.

Finally, students with total scores higher than 80.01% exhibited strong performance across all question sets. They were able to correctly solve about 3/4 of the integrated-atypical questions, indicating a profound understanding of the central idea. Therefore, these students are classified as having achieved an expertlike level of knowledge integration.

The score distribution across different question sets reveals the general progression of students' knowledge integration. Building upon the differences in students' performance illustrated in Fig. 4, a scale of 0-40-60-80-100 (in percentage) is identified to determine the different levels of knowledge integration as summarized in Table III. Here, it is also noted that the assessment of knowledge integration is population dependent. The scale provided here serves as a reasonable approximation based on the data in this study and may not be directly extended to examine other contexts and populations.

Based on the scale of students' knowledge integration levels, a comparison was made between high school students and college students, as depicted in Fig. 5. The analysis reveals that the majority of both high school and college students fall within the intermediate levels of knowledge integration. However, high school students predominantly occupy the lower-intermediate level, whereas college students are mostly concentrated at the upper-intermediate level. Furthermore, there are more high school students at the novice level compared to the college students, while the number of high school students at the

TABLE III. Summary of different question sets scores for each knowledge integration level for all students. Standard errors are given in brackets. The mean scores and standard errors are in percentage scale. The $p$ values reflect the significance of the one-way ANOVA analysis for the differences among mean scores of the novice, intermediate, and expertlike students in different question sets.

| Level | | Score range | $N$ | Single-link typical | Multilink atypical | Integrated-link atypical |
|---|---|---|---|---|---|---|
| Novice | | 0.00–40.00 | 76 | 49.08(1.35) | 28.13(1.54) | 13.01(1.18) |
| Intermediate | Lower intermediate | 40.01–60.00 | 194 | 71.60(0.86) | 46.07(1.12) | 31.73(0.91) |
| | Upper intermediate | 60.01–80.00 | 181 | 86.24(0.71) | 70.86(1.08) | 48.56(1.19) |
| Expertlike | | 80.01–100.00 | 41 | 95.37(0.93) | 87.80(1.41) | 71.00(1.86) |
| $p$ value | | | | $< 0.001$ | $< 0.001$ | $< 0.001$ |

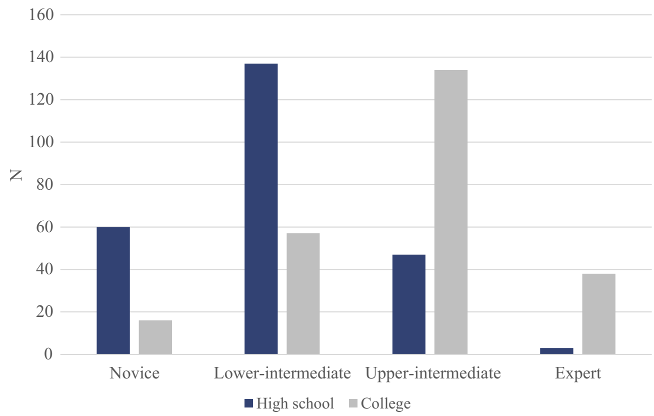


FIG. 5. Comparison the number of students at each level of knowledge integration among high school students and college students.

expertlike level is almost nonexisting. The results suggest that college students appear to have developed a deeper understanding of the central idea and the related concepts, which can be attributed to the emphasis in the college courses on the different electric potential at different points in galvanic cell circuits, rather than directly equating the magnitude of electrode potential of electrode materials with the differences in oxidation-reduction property of substances, which are typical instructional strategies in high school chemistry curricula. However, it is worth noting that only a small proportion of college students have achieved expertlike understanding, which warrants further research on more effective instruction for promoting knowledge integration.

In conclusion, the question design incorporating different link types and contextual saliency appears to be an effective strategy for evaluating students' knowledge integration levels. Students are categorized into four levels of knowledge integration based on their performances on different question sets.

### B. Qualitative research on students' reasoning pathways

To gain a deeper understanding of the reasoning pathways used by students at different levels of knowledge integration, think-out-loud interviews were conducted with students from each level to further validate the knowledge integration levels determined by the assessment scores. The students for interviews were selected by a research coordinator who did not participate in the interview study and data analysis. Students who completed the post-test were categorized into four knowledge integration levels based on their post-test scores. Then students in each level were randomly inquired for follow-up interviews, which generated a pool of 25 students, including 11 high school students and 14 college students. Among these 25 students, their knowledge integration levels categorized based on their total scores include six novices, nine lower intermediates, seven upper intermediates, and three expertlikes. The main objective of these interviews was to gain insights into students' misconceptions, reasoning pathways, and knowledge structures at different levels of knowledge integration. Thus, students at different knowledge integration levels were treated equally regardless if they were from high school or college.

For the analysis of the interviews, a codebook was developed to code students' answers and reasoning on each question into a corresponding knowledge integration level. This codebook is presented in Supplemental Material [62] and is developed based on the reasoning pathways of students at different levels of knowledge integration, which are modeled in Secs. IIB and IIC. During the interviews, students answered the same questions from the quantitative assessment and explained their reasoning throughout the process in think-aloud mode. Their responses were recorded, transcribed, and compared with the reasoning pathways outlined in the codebook to identify knowledge integration levels.

During the coding process, two researchers, who were not aware of the student's assessment scores to avoid confirmation biases, independently analyzed students' responses based on the codebook to determine each student's knowledge integration level. The common consistency between the two researchers was 0.90 with a Cohen's Kappa at 0.81, indicating that the independent coding results of the two researchers had good interrater reliability [59]. Any discrepancies were further discussed with a faculty to reach consensus. The codebook and coding process are discussed in Supplemental Material with the final outcomes listed in Table S4 [62]. The results revealed a 0.84 agreement between knowledge integration levels determined through the interviews and through the quantitative GCKIA scores, and the moderately high consistency between the quantitative and qualitative outcomes indicates a solid validity of the knowledge integration levels determined through the quantitative assessment [59]. Specifically, all six novices displayed consistent knowledge integration levels in both the interviews and the quantitative study; of the nine lower intermediates, seven showed consistent levels; of the seven upper intermediates, five had consistent levels; and all three expertlikes exhibited consistent knowledge integration levels in both settings. The discrepancies in the knowledge integration levels for some students were mainly because some students were able to correctly answer a few multilink atypical questions and integrated-link atypical questions on the GCKIA through guessing. However, their lack of understanding of the central idea and some intermediate concepts had led to incorrect reasoning in their explanations of the same questions during the interviews. Examples of students' reasoning identified through the interviews are discussed below to demonstrate the nuances of students' reasoning pathways.

#### *1. Novice level*

As shown from students' answers and explanations exemplified below, students at the novice level were only able to correctly answer a small fraction of the single-link typical questions. Interview results showed that most students were at the novice level. They solved problems by matching the surface features of the questions with memorized terms, laws, and equations, or by referencing similar questions that they had previously encountered, which are demonstrated with the students' excerpts shown below:

Student-6: *(Answered C to Q18. The correct answer is B) All four options of this question are correct. As taught by the teacher, at the negative electrode of the hydrogen-oxygen fuel cell,* $H_2$ *loses electrons to generate* $H^+$*, while at the positive electrode,* $O_2$ *accepts electrons to generate* $OH^-$*. Therefore, I randomly selected one option.*

Student-11: *(Answered D to Q21. The correct answer is B) The reactivity series of metals suggests that Cu has stronger reducibility than Ag, resulting in Cu being precipitated first.*

Student-24: *(Answered D to Q16. The correct answer is A) The reactivity series of metals suggests that Zn has stronger reducibility than Fe, leading to Zn being preferentially oxidized and serving as the negative electrode. Meanwhile,* $Cu^{2+}$ *at the positive electrode undergoes a reduction reaction, hence the positive electrode material is Cu.*

Like many typical novice students, these three students primarily relied on recalling patterns and formulas based on context variables and surface features to solve problems. For instance, *student 6* utilized memorized hydrogen-oxygen fuel cell reaction equations to solve problems but overlooked the fact that in alkaline conditions, $H^+$ and $OH^-$ cannot coexist. It is evident that students make superficial connections between context variables and potential outcomes, without connecting related concepts in a reasoning pathway that is scientifically valid. Additionally, *student 11* recalled the reactivity series of metals but failed to consider that the easier a metal can lose electrons, the weaker the metal's ion form can gain electrons. As a result, there is an incorrect decision on which metal ion should gain electrons.

Similarly, although *student 24* correctly recalled the reactivity series of metals for Q16, there appeared to be confusion regarding the concepts of electrode reactants and electrode materials. The student failed to grasp that the electrode material serves as a conduit for electron or ion transfer in a galvanic cell, while the electrode reactant undergoes oxidation or reduction in the reaction. This confusion might have stemmed from the slight differences in context between the question and the memorized case, as familiar questions typically inquire about the electrode reactant, leading to difficulty in correctly applying the memorized answer to the question presented.

The reasoning pathways of the aforementioned three students support the viewpoints articulated in the conceptual framework of galvanic cell, which demonstrated that the knowledge structures of novices are fragmented with limited connections and that they primarily rely on surface features as clues to recall memorized terms, laws, and equations in problem solving.

#### *2. Lower-intermediate level*

Students at this level can successfully solve many single-link typical questions, for which they can also provide more explanations by considering the oxidation-reduction property of substances. For instance, when solving Q25 and Q21, the following two students were able to determine the positive and negative electrodes and the corresponding chemical reactions in a galvanic cell based on the oxidation-reduction property of different substances.

Student 25: *(Answered C to Q25. The correct answer is C) When connecting a and b with wires and immersing them in the electrolyte solution, a is less susceptible to corrosion than b, indicating that the reducibility is* $b > a$*. When b and d are placed in hydrochloric acid of equal concentration, d reacts more vigorously than b, suggesting that the reducibility is* $d > b$*. When Cu is immersed in the salt solution of a, there is no noticeable change, indicating that the reducibility is* $a > Cu$*. If Cu is immersed in a salt solution of c, metal c will precipitate, indicating that the reducibility is* $Cu > c$*. Consequently, it can be deduced that the reducibility is* $a > Cu > c$*. Therefore, the reducibility changes from strong to weak:* $d > b > a > c$*.*

Student 4: *(Answered B to Q21. The correct answer is B) The reducibility of Zn is stronger than that of the inert electrode Pt, so the Zn electrode is the negative electrode. Zn loses electrons and becomes* $Zn^{2+}$*, causing the mass of the left electrode to decrease. On the right electrode, the reducibility of Cu is stronger than that of Ag, so the oxidizability of* $Cu^{2+}$ *is weaker than that of* $Ag^+$*. Therefore,* $Ag^+$ *first receives electrons to generate Ag. After all* $Ag^+$ *ions in the solution receive electrons and become Ag,* $Cu^{2+}$ *then receives electrons to generate Cu. Therefore, option B is chosen.*

Nevertheless, students at this level appeared to focus more on the application of the oxidation-reduction property of substances. Consequently, they encountered difficulties when solving multilink atypical questions. They struggled to identify the positive and negative electrodes based on the standard electrode potential, nor can they compare the electric potential at different points between positive and negative electrodes of the galvanic cell. Students who exhibited this behavior had thoughts similar to the interview excerpts shown below:

Student 5: *(Answered A to Q8. The correct answer is D) According to oxidation-reduction property, Fe is designated as the negative electrode, while graphite is identified as the positive electrode. Initially, as the reaction commences, there is no current flow, resulting in the absence of any potential difference. As the reaction progresses, electrons flow from Fe to graphite, causing a continuous decrease in the potential of graphite and a continuous increase in the potential of Fe.*

Student 18: *(Answered B to Q20. The correct answer is D) The standard electrode potentials of the two electrodes are added together. If the result is greater than 0, the galvanic cell can be formed. The electrode with the larger value is designated as the positive electrode, while the one with the smaller value is designated as the negative electrode.*

Student 1: *(Answered B to Q20. The correct answer is D) The positive and negative signs are meaningless because potential is a scalar quantity. A large absolute value of standard electrode potential indicates a higher potential, therefore it is used as the positive electrode. Conversely, a smaller absolute value of standard electrode potential indicates a lower potential, hence it is used as the negative electrode.*

Based on the interview results, one can see that neither *student 18* nor *student 1* had a correct understanding of the concept of standard electrode potential. Furthermore, *student 18* did not recognize that the electromotive force of a cell is determined by subtracting the negative electrode potential from the positive electrode potential. In addition, *student 5* lacked an understanding that a galvanic cell reaction can proceed spontaneously only when its electromotive force is greater than 0.

In summary, students at the lower-intermediate level demonstrated proficiency in solving problems based on the oxidation-reduction property of substances. However, they lacked an understanding of the standard electrode potential and the differing electric potential at different points in the external circuit, which prevented them from determining the positive and negative electrodes effectively. Consequently, they were only capable of solving single-link typical questions and were unable to correctly answer multilink atypical and integrated-link atypical questions. These students' reasoning corroborates with the lower-intermediate level of knowledge integration determined based on their total scores.

### *3. Upper-intermediate level*

Upper-intermediate level students outperformed lower-intermediate level students in solving multilink atypical questions. They demonstrated a stronger ability to establish connections with related concepts compared to lower-intermediate students. As shown by the interview excerpts below, students at this level correctly understood the concept of standard electrode potential and were able to analyze the electric potential at different points in the external circuit to determine the positive and negative electrodes.

Student 12: *(Answered D to Q20. The correct answer is D) In a galvanic cell, the reaction can proceed spontaneously only when the electromotive force is greater than 0. Therefore, the positive standard electrode potential must exceed the negative standard electrode potential. Considering that* $\varphi^{\Theta}(T^{2+}|T)$ *is less than* $\varphi^{\Theta}(X^{2+}|X)$*, when half-cell T is connected to half-cell X, half-cell X becomes the positive electrode.*

Student 17: *(Answered D to Q3. The correct answer is D) When A and B are combined to form a galvanic cell, and B dissolves, this indicates that B is the negative electrode. This implies that B has a stronger electron losing capability than A. Subsequently, when B is paired with C to form another galvanic cell, B's electrode potential is lower than C's. In external circuits, the electrode with lower potential is the negative electrode. Thus, B is the negative electrode again, indicating that B's ability to lose electrons is stronger than C's. In summary, B's ability to lose electrons is stronger than both A and C. Therefore, it is impossible to ascertain which would be the positive or negative electrode when A and C form a galvanic cell.*

However, these students have yet to develop an expert-level knowledge structure, lacking a good understanding of the central idea, which is exemplified by the following three students.

Student 13: *(Answered B to Q13. The correct answer is A) Aren't the standard electrode potentials of Cu and Fe fixed? Hence, the potential difference between Cu and the solution and Fe and the* *solution will remain unchanged before and after the reaction.*

Student 21: *(Answered B to Q13. The correct answer is A) The potential difference exists between the Fe rod and Cu rod, while the potential difference between the Fe rod and the solution and the Cu rod and the solution is always 0.*

Student 22: *(Answered D to Q4. The correct answer is B) Mg has stronger reducibility than Ag, so Mg is the negative electrode and Ag is the positive electrode, meaning Ag's potential is definitely higher than that of Mg. Furthermore, since both d and c are in the solution, the solution should be a uniform and stable equipotential body, so* $\varphi_d = \varphi_c$.

As shown by the interviews, students at the upper-intermediate level demonstrated improved comprehension of the differing electric potential at different points in the external circuit. However, they still lack a profound understanding of the causal mechanism for the potential differences across different contact interfaces between materials and/or solutions. As outlined in the conceptual framework of the galvanic cell, their approaches to solve galvanic cell problems still moderately rely on memorization, without developing a fundamental understanding of the causal mechanism (i.e., the central idea) underlying the various galvanic cell reactions.

#### 4. Expertlike level

As an indicator of achieving a conceptual understanding akin to that of experts, students at the expertlike demonstrated a profound comprehension of the causal mechanism for the generation of potential differences across contact interfaces between different materials and solutions. The responses provided by the following two students illustrated their grasp of the central idea.

Student 20: *(Answered B to Q1. The correct answer is B) The potential of the solution is 0. If the Fe rod is inserted into the water, some* $Fe^{2+}$ *ions will dissolve from the rod, while the electrons will remain on the inner surface of the Fe rod. Meanwhile,* $Fe^{2+}$ *ions will surround the outer surface of the Fe rod, eventually reaching an equilibrium state. Therefore, when the Fe rod is inserted into the water, its electric potential cannot be 0.*

Student 23: *(Answered B to Q4. The correct answer is B) Mg is more likely to lose electrons than Ag, resulting in the inner side of the Mg electrode being negatively charged and the outer side being positively charged. Conversely, the inner side of the Ag electrode is positively charged, while the outer side is negatively charged, thus leading to* $\varphi_c < \varphi_b$. *Regarding the question of why* $\varphi_c$ *is not equal to* $\varphi_d$, *it is because there are numerous positive charges surrounding the negative electrode and numerous negative charges surrounding the positive electrode. Since there is no salt bridge present to balance these charges, it is necessary to have* $\varphi_c > \varphi_d$.

Equipped with a good understanding of the central idea, expertlike students can adeptly navigate through different problem contexts to apply the central idea in problem solving. For instance, both Q19 and Q26 are integrated-link atypical questions, which require using the central idea to discern that the outer surface of the Fe electrode becomes positively charged while that of the Cu electrode becomes negatively charged based on the electron gaining or losing capabilities of the Fe electrode and the Cu electrode. In Q19, if the circuit is disconnected after a period of reaction without removing the salt bridge, the $K^+$ and $Cl^-$ ions in the salt bridge (KCl saturated solution) will naturally diffuse toward the positive and negative electrodes due to the disparate charge distributions in the two half-cells. Eventually, both half-cells reach electrical neutrality, resulting in $\varphi_a = \varphi_b$. In contrast, in Q26, if the salt bridge is removed without disconnecting the external circuit, the balancing effect of the salt bridge on the charges will be absent. Consequently, the solution in the negative electrode half-cell becomes negatively charged, while the solution in the positive electrode half-cell becomes positively charged, leading to $\varphi_a > \varphi_b$. The interview excerpts from the following two students demonstrate that they were able to engage in the expected reasoning similar to that of experts.

Student 23: *(Answered B to Q19. The correct answer is B) Fe is more likely to lose electrons, resulting in the side of the Fe electrode close to the solution being positively charged, while the side of the Cu electrode close to the solution becomes negatively charged. However, when a salt bridge is present, the negative charge within the bridge migrates towards the negative electrode, while the positive charge moves towards the positive electrode. Consequently, an equipotential body is established, leading to* $\varphi_a = \varphi_b$.

Student 15: *(Answered C to Q26. The correct answer is C) The primary reason lies in the electric double layer phenomenon. Due to Fe's stronger tendency to lose electrons, the inner surface of the electrode gathers negative charges while the outer surface accumulates positive charges, with Cu exhibiting the opposite behavior. In the absence of a salt*

> *bridge, the electrons cannot achieve equilibrium, leading to an imbalance in charge and resulting in $\varphi_a > \varphi_b$.*

According to the interview, it can be seen that these students had constructed reasoning pathways similar to those of experts, which enabled them to correctly answer integrated-link atypical questions. In solving problems, these students started by analyzing the charge distribution surrounding the electrodes based on the different electron gaining or losing capabilities of the electrodes. Subsequently, they leveraged this analysis to compare the electric potential between the internal and external circuits. Based on this knowledge structure, they were able to solve most integrated-link atypical questions, demonstrating a good understanding of the central idea and an expertlike level of knowledge integration.

In summary, the interviews from students at different knowledge integration levels determined by their total scores provided qualitative evidence of student reasoning and problem-solving approaches, which correspond well with the expected student behaviors defined in the knowledge integration framework. The results confirm that the GCKIA test is effective in probing features of knowledge integration in learning galvanic cell and that the scale of 0-40-60-80-100 (in percentage) is valid in categorizing students into four levels of knowledge integration.

## IV. CONCLUSION AND DISCUSSION

Achieving deep conceptual understanding through knowledge integration has been widely emphasized in science education. This study takes an interdisciplinary approach to knowledge integration, by introducing physics concepts as the central idea for developing a conceptual framework of galvanic cell to model student learning in the knowledge integration perspective. The conceptual framework is utilized to guide the development of an assessment instrument (GCKIA) for measuring students' levels of knowledge integration in learning galvanic cell. Building off the existing research on knowledge integration, the GCKIA was designed with three question sets in varied contextual saliency and knowledge connectedness including single-link typical, multilink atypical, and integrated-link atypical. Data were collected using GCKIA with Chinese high school and college students along with interviews with a subset of the students. The analysis of assessment data and interview results reveal that the assessment is effective in probing students' understanding of the central idea, which aids the categorization of students into four distinctive levels of knowledge integration including novice, lower-intermediate, upper-intermediate, and expertlike.

Novice students were only able to solve a few single-link typical questions and failed on most multilink atypical and integrated-link atypical questions. Their knowledge structures appeared fragmented and lacked relevance to the central idea. Students at this level exhibited a strong context dependence when solving problems, relying on matching contextual features with memorized terms, laws, and equations. Interviews revealed that while novice students might memorize some familiar problem contexts and corresponding solutions, they typically lacked even a basic understanding of the causal mechanism (i.e., the central idea) underlying the memorized fragments.

Intermediate students can be further categorized into two groups. Students at the lower-intermediate level demonstrated proficiency in solving single-link typical questions but still lacked a basic understanding of the central idea, leading to failure in more than half of the multilink atypical questions and most integrated-link atypical questions. On the other hand, students at the upper-intermediate level, in addition to being able to solve single-link typical questions, can also successfully solve a large fraction of multilink atypical questions but still have difficulties with many integrated-link atypical questions due to an incomplete understanding of the central idea. Specifically, these students appeared to have grasped the concept of standard electrode potential and could solve problems based on the differing electric potential at different points in the external circuit. However, they had yet to fully integrate this understanding with the central idea and lacked a deep comprehension of the causal mechanism underlying the many chemical reactions in galvanic cell reactions. Consequently, their knowledge structures remain locally linked between contextual features and intermediate procedures, rather than being fully connected to the central idea. Despite their progress, these students still exhibited some level of reliance on memorized procedures and rules in problem solving.

Expertlike students demonstrated the ability to solve all types of problems, especially the integrated-link atypical questions. Their reasoning in problem solving demonstrated a comprehensive understanding of the central idea, which forms the foundation of an integrated knowledge structure. More importantly, they exhibited advanced reasoning using the central idea, which empowered them to navigate through unfamiliar problems without being constrained by contextual features.

The assessment results of high school and college students in China also revealed that the majority of these students are at the intermediate levels of knowledge integration. Only a small portion of college students and a minimal number of high school students achieved the expertlike level. The results suggest that most Chinese students developed relatively fragmented knowledge structures in their learning through existing curricula on galvanic cell.

The results of this study also revealed the importance of providing and emphasizing explanatory mechanisms for students in learning, which can help them integrate the many complex phenomena and processes into a coherent knowledge system including not only factual content but also meaningful reasoning to connect the different knowledge components. Following the interdisciplinary knowledge integration approach, education intervention studies are warranted and are examined in our current and future studies.

Although this study has yielded encouraging results, there are limitations that warrant further investigation in future research. The study sample was limited, including students from three middle-ranked high schools in the southwestern and southeastern provinces of China, and undergraduates from universities ranked within the top 150 in China. Future research should broaden the sample to include students from high schools and universities of different levels. Another limitation arising from the limited sample is that only a small number of high-performing students were identified, which restricted the analysis of expert-level students. Future research would benefit from including different advanced student groups, such as graduate students specializing in electrochemistry. This inclusion would allow for a more comprehensive examination of the development process of interdisciplinary knowledge integration in galvanic cell.

Because interdisciplinary insights have often provided essential explanatory frameworks for understanding specific fields, it is crucial to conduct interdisciplinary knowledge integration to help students form a comprehensive knowledge structure. This paper introduces the important role of the conceptual framework model in promoting knowledge integration of the interdisciplinary topic of galvanic cell and discusses the assessment study carried out under the guidance of this framework. This study can offer practical benefits in multiple ways. For example, it can provide policymakers with a reference on how to break down disciplinary boundaries and build more explanatory knowledge frameworks. It also offers teachers and researchers new, reliable tools for the assessment of interdisciplinary knowledge integration instructional designs and teaching practices. Policymakers can incorporate the galvanic cell conceptual framework into curriculum standards and textbooks to promote deep learning from a curriculum design perspective. Researchers and teachers can design galvanic cell instruction based on the conceptual framework and use the GCKIA to evaluate changes in students' knowledge structures before and after instructional intervention, thereby helping students form a more comprehensive knowledge structure through teaching and assessment.

Additionally, the results validate the efficacy of employing the conceptual framework for assessment and instruction, showcasing its ability to enhance students' knowledge integration and deepen their understanding. This study not only introduces a new exemplary case based on previous research but also expands the feasibility and practical value of the conceptual framework model in interdisciplinary education, extending the boundary of previous applications that were conducted solely in the field of physics education.

## ACKNOWLEDGMENTS

This research is supported in part by the China Scholarship Council (No. 202206320152), the Teaching Reform Project of Zhejiang University and Yuhang High School Cooperation (No. 2022-SKY-503103-0013), and the National Science Foundation Grant No. DUE-2043817.

Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agencies.

[1] C. S. Wagner *et al.*, Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature, J. Informetr. **5**, 14 (2011).

[2] S. Wang, J. Mao, Y. Cao, and G. Li, Integrated knowledge content in an interdisciplinary field: Identification, classification, and application, Scientometrics **127**, 6581 (2022).

[3] F. Morillo, M. Bordons, and I. Gómez, Interdisciplinarity in science: A tentative typology of disciplines and research areas, J. Am. Soc. Inf. Sci. Technol. **54**, 1237 (2003).

[4] I. Rafols and M. Meyer, Diversity and network coherence as indicators of interdisciplinarity: Case studies in bionanoscience, Scientometrics **82**, 263 (2010).

[5] B. S. Bloom, M. D. Engelhart, E. J. Furst, W. H. Hill, and D. R. Krathwohl, *Taxonomy of Educational Objetives: The Classification of Educational Goals: Handbook I: Cognitive Domain* (D. Mckay, New York, 1956).

[6] J. Conklin, *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives Complete Edition, Abridged Edition* (2005).

[7] O. Gulacar, B. Vernoy, E. Tran, A. Wu, Emily Zhang Huie, Efrain Vasquez Santos, Anish Wadhwa, Risa Sathe, and Alexandra Milkey, Investigating differences in experts' chemistry knowledge structures and comparing them to those of general chemistry students, J. Chem. Educ. **99**, 2950 (2022).

[8] G. Papageorgiou, A. Markos, and N. Zarkadis, Students' representations of the atomic structure–the effect of some

individual differences in particular task contexts, Chem. Educ. Res. Pract. **17**, 209 (2016).

[9] C.-Y. Wang and L. H. Barrow, Exploring conceptual frameworks of models of atomic structures and periodic variations, chemical bonding, and molecular shape and polarity: A comparison of undergraduate general chemistry students with high and low levels of content knowledge, Chem. Educ. Res. Pract. **14**, 130 (2013).

[10] L. Bao and E. F. Redish, Model analysis: Representing and assessing the dynamics of student learning, Phys. Rev. ST Phys. Educ. Res. **2**, 010103 (2006).

[11] R. M. Heyworth, Procedural and conceptual knowledge of expert and novice students for the solving of a basic problem in chemistry, Int. J. Sci. Educ. **21**, 195 (1999).

[12] J. Bransford, A. Brown, and R. Cocking, *How Experts Differ from Novices in How People Learn* (National Academy Press, Washington DC, 1999).

[13] C. Lu, Y. Liu, S. Xu, S. Zhou, H. Mei, X. Zhang, L. Yang, and L. Bao, Conceptual framework assessment of knowledge integration in student learning of measurement uncertainty, Phys. Rev. Phys. Educ. Res. **19**, 020145 (2023).

[14] L. Bao and J. C. Fritchman, Knowledge integration in student learning of Newton's third law: Addressing the action-reaction language and the implied causality, Phys. Rev. Phys. Educ. Res. **17**, 020116 (2021).

[15] T. Vogt, The value of vague ideas in the development of the periodic system of chemical elements, Synthese **199**, 10587 (2021).

[16] L. A. Elliott, E. Sippola, and J. Watkins, Modeling chemical reactions with the Gaussian gun, J. Chem. Educ. **96**, 100 (2018).

[17] M. M. Cooper and M. W. Klymkowsky, The trouble with chemical energy: Why understanding bond energies requires an interdisciplinary systems approach, CBE Life Sci. Educ. **12**, 306 (2013).

[18] D. Tong, J. Liu, Y. Sun, Q. Liu, X. Zhang, S. Pan, and L. Bao, Assessment of student knowledge integration in learning work and mechanical energy, Phys. Rev. Phys. Educ. Res. **19**, 010127 (2023).

[19] L. Xie, Q. Liu, H. Lu, Q. Wang, J. Han, X. Feng, and L. Bao, Student knowledge integration in learning mechanical wave propagation, Phys. Rev. Phys. Educ. Res. **17**, 020122 (2021).

[20] Z. Liu, S. Pan, X. Zhang, and L. Bao, Assessment of knowledge integration in student learning of simple electric circuits, Phys. Rev. Phys. Educ. Res. **18**, 020102 (2022).

[21] L. Bao and J. C. Fritchman, Knowledge integration in student learning of Newton's third law: Addressing the action-reaction language and the implied causality, Phys. Rev. Phys. Educ. Res. **17**, 020116 (2021).

[22] W. Xu, Q. Liu, K. Koenig, J. Fritchman, J. Han, S. Pan, and L. Bao, Assessment of knowledge integration in student learning of momentum, Phys. Rev. Phys. Educ. Res. **16**, 010130 (2020).

[23] R. Dai, J. C. Fritchman, Q. Liu, Y. Xiao, H. Yu, and L. Bao, Assessment of student understanding on light interference, Phys. Rev. Phys. Educ. Res. **15**, 020134 (2019).

[24] Y. Nie, Y. Xiao, J. C. Fritchman, Q. Liu, J. Han, J. Xiong, and L. Bao, Teaching towards knowledge integration in learning force and motion, Int. J. Sci. Educ. **41**, 2271 (2019).

[25] K. L. Turner, S. He, B. Marchegiani, S. Read, J. Blackburn, N. Miah, and M. Leketas, Around the world in electrochemistry: A review of the electrochemistry curriculum in high schools, J. Solid State Electrochem. **28**, 1361 (2024).

[26] S. Rahayu, D. F. Treagust, A. Chandrasegaran, M. Kita, and S. Ibnu, Assessment of electrochemical concepts: A comparative study involving senior high-school students in Indonesia and Japan, Res. Sci. Technol. Educ. **29**, 169 (2011).

[27] J. Wang and C. Zheng, *Chemistry: Principles of Chemical Reactions* (People's Education Press, Beijing, 2019), pp. 93–103.

[28] O. Keiichiro and M. Motoyuki, *Chemistry: Theoretical Edition* (Tokyo Shoseki Co., Ltd., Tokyo, 2023), pp. 112–119.

[29] P. J. Garnett and D. F. Treagust, Conceptual difficulties experienced by senior high school students of electrochemistry: Electrochemical (galvanic) and electrolytic cells, J. Res. Sci. Teach. **29**, 1079 (1992).

[30] H. Lu, Y. Jiang, and H. Bi, Development of a measurement instrument to assess students' proficiency levels regarding galvanic cells, Chem. Educ. Res. Pract. **21**, 655 (2020).

[31] T. Buthelezi, L. Dingrando, N. Hainen, C. Wistrom, and D. Zike, *Glencoe Chemistry: Matter & Change* (McGraw Hill Education, Houston, TX, 2017), pp. 708–727.

[32] L. Derry *et al.*, *Heinemann Chemistry 2* (Pearson Australia, Melbourne, Victoria, 2017), pp. 125–168.

[33] B. Acar and L. Tarhan, Effect of cooperative learning strategies on students' understanding of concepts in electrochemistry, Int. J. Sci. Math. Educ. **5**, 349 (2007).

[34] M. T. Ali, A. R. Woldu, and A. G. Yohannes, High school students' learning difficulties in electrochemistry: A mini review, Afr. J. Chem. Educ. **12**, 202 (2022), https://www.ajol.info/index.php/ajce/article/view/229361.

[35] A. R. Özkaya, M. Üce, H. Saricayir, and M. Sahin, Effectiveness of conceptual change-oriented teaching strategy to improve students' understanding of galvanic cells, J. Chem. Educ. **83**, 1719 (2006).

[36] A. R. Özkaya, Conceptual difficulties experienced by prospective teachers in electrochemistry: Half-cell potential, cell potential, and chemical, and electrochemical equilibrium in galvanic cells, J. Chem. Educ. **79**, 735 (2002).

[37] A. Yilmaz and S. Bayrakçeken, Determining of the prospective teachers' understadings of electrochemistry, Procedia **174**, 2831 (2015).

[38] M. J. Sanger and T. J. Greenbowe, Students' misconceptions in electrochemistry regarding current flow in electrolyte solutions and the salt bridge, J. Chem. Educ. **74**, 819 (1997).

[39] A. R. Şekerci and A. R. Erdem, The development and implementation of a two-tier multiple choice test related to the concepts of oxidation-reduction reactions and galvanic cells, J. Educ. Sci. Psychol. **12**, 34 (2022).

[40] P. J. Garnett and D. F. Treagust, Conceptual difficulties experienced by senior high school students of electrochemistry: Electric circuits and oxidation-reduction equations, J. Res. Sci. Teach. **29**, 121 (1992).

[41] M. J. Sanger and T. J. Greenbowe, Common student misconceptions in electrochemistry: Galvanic, electrolytic, and concentration cells, J. Res. Sci. Teach. **34**, 377 (1997).

[42] M. J. Sanger and T. J. Greenbowe, An analysis of college chemistry textbooks as sources of misconceptions and errors in electrochemistry, J. Chem. Educ. **76**, 853 (1999).

[43] L. T. Tien and K. Osman, Misconceptions in electrochemistry: How do pedagogical agents help?, in *Overcoming Students' Misconceptions in Science: Strategies and Perspectives from Malaysia*, edited by M. Karpudewan, A. N. Md Zain, and A. L. Chandrasegaran (Springer Singapore, Singapore, 2017), pp. 91–110.

[44] A. S. L. Loh, R. Subramaniam, and K. C. D. Tan, Exploring students' understanding of electrochemical cells using an enhanced two-tier diagnostic instrument, Res. Sci. Technol. Educ. **32**, 229 (2014).

[45] P. J. Garnett, P. J. Garnett, and M. W. Hackling, Students' alternative conceptions in chemistry: A review of research and implications for teaching and learning, Stud. Sci. Educ. **25**, 69 (1995).

[46] F. Karsli, Can freshman science student teachers' alternative conceptions of 'electrochemical cells' be fully diminished?," Asian J. Chem. **24**, 485 (2012), https://asianpubs.org/index.php/ajchem/article/view/8584/8572.

[47] B. Acar Sesen and L. Tarhan, Inquiry-based laboratory activities in electrochemistry: High school students' achievements and attitudes, Res. Sci. Educ. **43**, 413 (2013).

[48] D. P. Rosenthal and M. J. Sanger, Student misinterpretations and misconceptions based on their explanations of two computer animations of varying complexity depicting the same oxidation–reduction reaction, Chem. Educ. Res. Pract. **13**, 471 (2012).

[49] F. Rogers, P. A. Huddle, and M. D. White, Using a teaching model to correct known misconceptions in electrochemistry, J. Chem. Educ. **77**, 104 (2000).

[50] J. D. Novak and A. J. Cañas, The origins of the concept mapping tool and the continuing evolution of the tool, Inf. Vis. **5**, 175 (2006).

[51] S. Rahayu, D. F. Treagust, and A. Chandrasegaran, High school and preservice chemistry teacher education students' understanding of voltaic and electrolytic cell concepts: Evidence of consistent learning difficulties across years, Int. J. Sci. Math. Educ. **20**, 1859 (2022).

[52] M. Niaz, Facilitating conceptual change in students' understanding of electrochemistry, Int. J. Sci. Educ. **24**, 425 (2002).

[53] H. J. Schmidt, A. Marohn, and A. G. Harrison, Factors that prevent learning in electrochemistry, J. Res. Sci. Teach. **44**, 258 (2007).

[54] K. Osman and T. T. Lee, Impact of interactive multimedia module with pedagogical agents on students'understanding and motivation in the learning of electrochemistry, Int. J. Sci. Math. Educ. **12**, 395 (2014).

[55] O. L. Liu, H. S. Lee, and M. C. Linn, Measuring knowledge integration: Validation of four-year assessments, J. Res. Sci. Teach. **48**, 1079 (2011).

[56] M. C. Linn, The knowledge integration perspective on learning and instruction, in *The Cambridge Handbook of the Learning Sciences*, edited by R. K. Sawyer (Cambridge University Press, New York, 2006), pp. 243–264, 10.1017/CBO9780511816833.016.

[57] J. L. Chiu and M. C. Linn, Supporting knowledge integration in chemistry with a visualization-enhanced inquiry unit, J. Sci. Educ. Technol. **23**, 37 (2014).

[58] M. C. Linn, H.-S. Lee, R. Tinker, F. Husic, and J. L. Chiu, Teaching and assessing knowledge integration in science, Science **313**, 1049 (2006).

[59] M. L. McHugh, Interrater reliability: The kappa statistic, Biochem. Med. **22**, 276 (2012), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3900052/.

[60] R. F. DeVellis and C. T. Thorpe, *Scale Development: Theory and Applications* (Sage Publications, Thousand Oaks, CA, 2021).

[61] N. Bedduside, A. Hadis, A. A. Jalal, and D. Patongai, Inovation of biology learning through the development of authentics assessment based on scientific literacy for student of senior high school, J. Phys. Conf. Ser. **1899**, 012142 (2021).

[62] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevPhysEducRes.20.020129 for galvanic cell knowledge integration assessment, statistical analysis of the assessment test, and codebook of galvanic cell knowledge integration assessment.

[63] J. P. Meyer and E. Hailey, A study of Rasch, partial credit, and rating scale model parameter recovery in WINSTEPS and jMetrik, J. Appl. Meas. **13**, 248 (2012), http://jampress.org/abst2012.htm.

[64] W.-C. Wang and C.-T. Chen, Item parameter recovery, standard error estimates, and fit statistics of the WINSTEPS program for the family of Rasch models, Educ. Psychol. Meas. **65**, 376 (2005).

[65] K. B. Christensen, G. Makransky, and M. Horton, Critical values for Yen's Q 3: Identification of local dependence in the Rasch model using residual correlations, Appl. Psychol. Meas. **41**, 178 (2017).

[66] W. J. Boone, Rasch analysis for instrument development: Why, when, and how?, CBE Life Sci. Educ. **15**, rm4 (2016).

[67] F. Franchignoni, G. Mora, A. Giordano, P. Volanti, and A. Chiò, Evidence of multidimensionality in the ALSFRS-R Scale: A critical appraisal on its measurement properties using Rasch analysis, J. Neurol., Neurosurg., Psychiatry **84**, 1340 (2013).

[68] O. N. Medvedev, R. J. Siegert, P. Kersten, and C. U. Krägeloh, Rasch analysis of the Kentucky inventory of mindfulness skills, Mindfulness **7**, 466 (2016).

[69] T. Z. H. T. Petra and M. J. Ab Aziz, Investigating reliability and validity of student performance assessment in Higher Education using Rasch Model, J. Phys. Conf. Ser. **1529**, 042088 (2020).