Journal of Research in Science Teaching

WILEY

Check for updates

RESEARCH ARTICLE OPEN ACCESS

Building Connections to Teach the Nature of Science: An Experienced Science Teacher's Formative Assessment Practices in a High School Classroom

Wonyong Park[1] | Sibel Erduran[2] | Judith Hillier[2]

[1]Southampton Education School, University of Southampton, Southampton, UK | [2]Department of Education, University of Oxford, Oxford, UK

Correspondence: Wonyong Park (w.park@soton.ac.uk)

Received: 17 February 2025 | **Revised:** 24 October 2025 | **Accepted:** 13 November 2025

Keywords: conversation analysis | formative assessment | nature of science | teacher practices

ABSTRACT

While understanding the nature of science (NOS) is recognized as a key component of science education because of its potential benefits in nurturing scientific literacy, there is currently limited knowledge on how NOS learning can be supported by teachers' ongoing, in-the-moment assessment in the classroom. This study explores the value of formative assessment as a vehicle to facilitate reflective instruction of the NOS, through building connections between NOS and other related components of science learning. We closely examined one experienced science teacher's lessons on the development of scientific knowledge in a 10th grade science classroom, along with the teacher's post-lesson reflections and classroom artifacts, with a focus on how the teacher planned and orchestrated whole-class discussions following student group presentations. We identify four distinct strategies for formative assessment the teacher enacted in an effort to create "teachable moments" for NOS learning: multicontextualizing NOS, mapping examples to NOS, focusing discussion on NOS, and exploring different aspects of NOS. Across the lessons, while some teacher moves prompted discursive exchanges, others revealed significant challenges and were often met with student silence. Our work uses formative assessment as a lens to reveal the complexities and in-the-moment challenges of NOS instruction, illustrating how even an experienced teacher can struggle to facilitate these conversations. We argue that NOS research should pay more attention to the potential of formative assessment and the need for professional development to support teachers with empirically grounded assessment tasks for use in NOS instruction.

1 | Introduction

The nature of science (NOS) has been advocated as a core component of the science curriculum to accomplish scientific literacy. Accordingly, there has been a great deal of interest among science educators in assessing learners' understanding of the NOS. It has been argued that for effective NOS teaching, it should be "intentionally planned for, taught and *assessed*" (Lederman et al. 2001, 137, italics added), suggesting that high-quality assessments should be considered an essential condition for the success of NOS instruction (Allchin 2011; Hanuscin et al. 2011). The NOS refers to "the epistemology of science, science as a way of knowing, and the values and beliefs inherent to scientific knowledge and its development" (Lederman 2007, 833). Recent studies that analyzed standards, curricula and curriculum guidelines suggest that NOS is recognized as a valuable component of science curriculum and standards documents in many jurisdictions (Erduran and Dagher 2014; Ferreira and Morais 2013; Olson 2018; Park, Wu, and Erduran 2020; Park, Yang, and Song 2020; Reinisch and Fricke 2022). Despite the widespread curricular inclusion of the NOS as a goal of science education, reviews of NOS literature across decades have identified that many students do not possess adequate knowledge of the NOS (Abd-El-Khalick and Lederman 2023; Deng et al. 2011; Lederman 2007). In addition, research has found that teachers' knowledge of the NOS is not automatically translated into classroom practice, and that teachers do not consider NOS to be an instructional outcome of equal status with that of "traditional" subject matter outcomes (Lederman 2007).

Empirical research on NOS instruction has generated a strong evidence base that NOS learning requires "explicit-reflective" instruction, as confirmed by a recent meta-analysis (Deng et al. 2025). Making a distinction between the "explicit" and "reflective" labels, Abd-El-Khalick (2013) argued that the former has curricular implications where NOS should be targeted as an explicit learning goal of the lesson (as opposed to a side effect of learning other things), while the latter is concerned with specific instructional approaches where the provision of structured learning opportunities are ensured. Importantly, reflective NOS teaching includes "providing students with opportunities to analyze their activities from within a NOS framework, map *connections* between these activities and those of scientists, and make conclusions about scientific epistemology" (Abd-El-Khalick and Akerson 2004, 792, italics added). Given the close relationships between NOS learning and the learning of scientific concepts, scientific practices, the history and philosophy of science (Khishfe 2023; McDonald 2010; Ryder and Leach 2008), teachers' capacity to facilitate connections between various levels and categories of knowledge would be essential to facilitate NOS learning. Similarly, an emerging body of research using epistemic network analysis methods have stressed the connections between NOS ideas as a core learning outcome (Mulvey et al. 2021; Peters-Burton et al. 2019).

Presently, the most often used instrument for assessing NOS is Views of Nature of Science (VNOS) and its variants (Lederman et al. 2002), which consist of open-ended questions about several key NOS aspects. Although such researcher-developed instruments have the potential to inform classroom instruction and assessment of the NOS, it should be noted that classroom assessment has aims that are distinct from those of standardized measurement for research purposes. As Allchin et al. (2014) pointed out, teachers often have other considerations for classroom assessment such as time constraints and the institutional context of accountability (Bell 2007; Sadler 2014). With the use of NOS assessment instruments, researchers are typically interested in collecting data to investigate patterns and trends, such as understanding the distribution of students' informed, mixed, and naïve NOS views, comparing these results across different groups, or tracing changes between pre- and post-interventions in diverse instructional settings (Abd-El-Khalick 2014). More recently, Abd-El-Khalick et al. (2024) developed a rubric to analyze and score responses to the VNOS, providing systematic guidance for qualitative interpretation and categorization, and explicitly scaffolding qualitative inferencing to standardize scoring and reduce both the burden and variability of VNOS analysis.

A further development in recent literature is the incorporation of multimodality in assessing NOS. For example, Cheung and Erduran (2025) developed an analytical framework that characterizes students' multimodal representations of the nature of scientific practices and methods which extend the previous emphasis in the literature on textual representations of NOS. These researchers applied the instrument to the analysis of 40 7th-grade students and examined their responses before and after engaging in explicit-reflective instruction focusing on NOS. Students' ideas changed after students engaged in NOS instruction that involved iterative compositions of multimodal representations of NOS. The investigation of multimodal meaning-making about NOS (Cheung et al. 2025) presents a new avenue for formative assessment of NOS not only in students' but also in teachers' learning of NOS (Barak et al. 2023; Erduran and Kaya 2018) in future studies.

While NOS researchers' aims may overlap with the aims of classroom assessment, it is unlikely that research instruments are sufficient to fulfill *teachers*' and *learners*' needs in the classroom. In the classroom setting, teachers are interested in facilitating learning by providing feedback to student responses and improving instruction (the *formative* use of assessment) and producing evidence of students' development at a particular time to make judgments about individual learners or the effectiveness of education (the *summative* use of assessment) (Black and Wiliam 2009). NOS research instruments have typically been employed in ways that align with summative rather than formative purposes, as they are often used to document patterns or shifts in students' NOS views across cohorts or interventions. Meanwhile, teachers can embed assessment in different forms of learning activities such as teacher–student interactions, peer interactions, group activities and portfolios, besides paper-and-pencil tests, to gauge students' NOS learning minute by minute, and day by day. Such differences necessitate considering the classroom assessment of the NOS as an activity that has distinct aims, methods and criteria for quality from the assessment through research instruments such as VNOS. This study aims to investigate one teacher's enactment of formative assessment in a 10th-grade lesson focused on the NOS, to explore how formative assessment can support NOS learning. Two research questions guided the study: (1) How did the teacher's formative assessment strategies create opportunities for NOS learning? (2) What challenges did the teacher encounter in using formative assessment to support NOS learning?

2 | Review of Literature

2.1 | Teachers' Use of Formative Assessment in Science Classroom

Teachers' interactions with students are instrumental in creating learning opportunities and engaging students in making sense of their science learning (Deverel-Rico and Furtak 2025; Mercer et al. 2009). One useful lens to analyze teacher–student interactions is formative assessment. The focus of formative assessment, or assessment *for* learning (as opposed to assessment *of* learning), is to support learning rather than report the achievements or satisfy accountability demands (Black 1998). Research and policy in science education have recognized the value of formative assessment and the important role of teachers' formative assessment skills in supporting science learning (Bell and Cowie 2001; Furtak 2023; Ruiz-Primo and Furtak 2007; National Research Council 2001). On what counts as formative assessment, Black and Wiliam (2009) argued that:

> Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited. (p. 9)

From this definition, it can be inferred that formative assessment aims for (a) the improvement of teaching as well as learning, through a cyclical and collaborative process, and (b) the inclusion of teachers, learners and peers as sources of learning. The improvement of teaching through formative assessment can occur over different time scales (Wiliam 2009): it can occur in seconds and minutes as the teacher responds to evidence; in days and weeks as the teacher designs and refines subsequent lessons; and in months and years as she prepares the same unit for the next year's cohort. Black and Wiliam also suggested that formative assessment is concerned with "the creation of, and capitalization upon, 'moments of contingency' in instruction for the purpose of the regulation of learning processes" (2009, 6). Furthermore, sociocultural perspectives on formative assessment highlight how knowledge is socially constructed through "assessment conversations" in which ideas are shared, negotiated, and articulated by teachers and students (Duschl and Gitomer 1997; Ruiz-Primo 2011). Such a discursive and sociocultural nature of formative assessment makes it a collaborative endeavor between teachers and students with the goal of improving learning, often involving issues of power (Cizek et al. 2019).

Several models have been proposed to capture the dynamic and situated nature of informal formative assessment (Bell and Cowie 2001). One influential framework is the ESRU (eliciting, student response, recognizing, and using information) model developed by Ruiz-Primo and Furtak (2007), which provides a systematic way to describe how teachers and students co-construct formative assessment moments in science classrooms. They applied the ESRU model to video analyses of science classrooms, demonstrating that productive formative assessment sequences are those in which teachers not only recognize the substance of students' ideas but also use them to guide subsequent instruction rather than simply evaluate correctness. Models such as the ESRU underscore that informal formative assessment is both epistemic and pedagogical: it requires teachers to discern the scientific meaning in students' utterances and to respond in ways that advance understanding.

Formative assessment ranges from spontaneous, in-the-moment interactions between students and teachers (i.e., informal formative assessment) to structured, pre-planned tasks and activities that help assess the current state of learning (i.e., formal formative assessment) (Shavelson et al. 2008). This study considers both *formal* and *informal* formative assessment by the teacher. Formal formative assessment takes place when a teacher plans for a specific set of tasks and/or questions that are targeted at the learning goal of the lesson in order to check for student understanding (Shavelson et al. 2008), and students produce outputs that can be used for assessing and possibly grading (Bell 2008). On the contrary, informal formative assessment is not wholly planned ahead of time but occurs spontaneously when the teacher notices a teachable moment during classroom conversation and seizes the opportunity to learn (Nieminen et al. 2016). These different types of formative assessments can be used to make science teaching more responsive to students' prior knowledge and experience (Empson and Jacobs 2008; Furtak, Thompson, et al. 2016; Furtak, Kiemer, et al. 2016) and close the achievement gap between high- and low-performing students (Yeh 2010). Given the importance of teachers' formative assessment in noticing and responding to individuals' learning needs and thereby providing equitable science learning for all learners (Furtak and Lee 2023; Kang and Furtak 2021), it is crucial to understand how teachers can create opportunities for learning through the use of various types of formative assessment and, subsequently, how to support teachers' abilities to do so. This way, both formal and informal formative assessment are inseparable from the process of teaching and learning.

2.2 | NOS Formative Assessment: An Underexplored Research Area

Research on NOS instruction has suggested that teaching the NOS requires knowledge in diverse domains throughout scientific content, history and philosophy of science, instructional strategies (Abd-El-Khalick 2013; Wahbeh and Abd-El-Khalick 2014); teachers' understandings of the NOS are necessary for teaching the NOS in classrooms but are not sufficient (Abd-El-khalick and Lederman 2000). There are reported constraints that teachers experience, arising from the complexity of students' cognitive and affective abilities (Abd-El-Khalick et al. 1998), lack of teachers' NOS-specific skills (Schwartz and Lederman 2002), insufficient educational resources (Wong et al. 2009), and teachers' attitudes and beliefs about the importance of NOS (Lederman 1999).

When it comes specifically to assessment, a few authors have suggested that teachers' knowledge of how to use classroom assessment to create NOS learning opportunities comprises a core part of pedagogical content knowledge for teaching the NOS (Hanuscin et al. 2020; Mesci et al. 2020; Schwartz and Lederman 2002). Yet, despite the significance of assessment in NOS learning and the need for understanding teachers' NOS instructional practices, there is currently limited knowledge about how formative assessment can help teachers to foster students' NOS learning in their classrooms (Allchin 2013; Hanuscin et al. 2011, 2020). Studies have found that teachers tend not to assess students' NOS understanding in their classrooms, formatively or summatively. In a study with elementary teachers on a 3-year professional development program, Hanuscin et al. (2011) reported that teachers mostly relied on using the instruments provided by researchers, but no attempts were made to use their own formative or summative assessment strategies to determine students' learning of the NOS. In secondary science, Abd-El-Khalick et al. (1998) reported that none of the 14 preservice secondary science teachers who participated in their study made any attempt to *assess* NOS learning, although most of them believed in the importance of *teaching* the NOS in schools. Such discrepancy between the perceived importance and assessment practice related to the NOS has been confirmed in another recent study of secondary science teachers in England (Brock and Taber 2019). Similarly, Ryder and Leach's (2008) analysis of teacher talk in high school classrooms suggested that while teachers were often able to elicit students' NOS understandings during whole-class discussions, they tended to end the discussions with non-evaluative comments ("fine, fine") without assessing students' understandings against the NOS learning goal or using it to create further learning opportunities.

In the present study, we focus on how NOS formative assessment can manifest in the specific context of high school science instruction. While there are some generic features of formative assessment that are common to many disciplines, assessment researchers have acknowledged the importance of "discipline-specific elements that comprise the formal and informal materials, collaborative processes, ways of knowing, and habits of mind particular to a content domain" (Cizek et al. 2019, 14). Given that NOS learning goals can vary from forming certain beliefs about the epistemic features of science to applying NOS ideas to specific historical and contemporary examples (Brock and Park 2024), and correspondingly require new pedagogical knowledge and skills for the teachers (Hanuscin et al. 2011; Mesci et al. 2020), formative assessment of NOS learning deserves separate attention from the general discussion of formative assessment in science education.

3 | Formative Assessment in Support of Explicit-Reflective NOS Instruction

We posit that formative assessment is central to facilitating NOS learning because it shares key features with explicit-reflective instruction. First, formative assessment creates a safe space for students to articulate and reflect on NOS ideas (McMillan 2010). Second, it is iterative, allowing students to revisit and refine their NOS understanding through ongoing feedback (Black and Wiliam 2009). Third, teachers can adapt instruction by using assessment data to clarify misconceptions and reinforce accurate understanding (Furtak, Thompson, et al. 2016; Furtak, Kiemer, et al. 2016). Fourth, it provides a lens for understanding sociocultural exchange through "assessment conversations" (Duschl and Gitomer 1997).

This section draws on Leahy et al.'s (2005) conceptualization of formative assessment that consists of five key instructional strategies. Table 1 is an adapted version of their framework to explicit-reflective NOS instruction, highlighting the convergences. The listed strategies result from crossing three processes (where the learner is going, where the learner is right now, and how to get there) (National Research Council 2001) with three groups of agents in the classroom (teacher, peer, learner). In the following, we discuss how each of these strategies can relate to NOS teaching based on the literature on explicit-reflective NOS instruction. In the study, we used this framework to interpret teachers' actions and conversations in a NOS classroom through the lens of formative assessment.

3.1 | Clarifying, Sharing, and Understanding Learning Intentions

Effective learning requires aligning assessment and instruction (American Educational Research Association et al. 2014; Martone and Sireci 2009). For formative assessment to work, teachers must share clear, transparent learning objectives and success criteria, often by making standards student-friendly (Leahy et al. 2005). This aligns with explicit NOS instruction, which directly targets a cognitive understanding of NOS (Khishfe and Abd-El-Khalick 2002). Importantly, teachers should also share a *justification* of NOS learning goals (Roberts et al. 1992), that is, tell students why they should learn about the NOS. Since students are often "unclear about the intended learning outcomes" for epistemic topics compared to content knowledge (Ryder and Leach 2008, 293), sharing the NOS objective at the outset and consistently reinforcing it is critical to guide the teaching and learning process.

3.2 | Eliciting Evidence of Learning

Formative assessment can provide a way to elicit and gauge information about students' current levels of achievement throughout the course of instruction. This function of formative assessment is particularly relevant to NOS learning, where students' understanding may be implicit or difficult to distinguish from their understanding of content knowledge or inquiry skills. In the context of reflective NOS instruction, elicitation includes bringing students' "tacit, fragmented and inarticulate" NOS understandings to the forefront and examining and revising them (Abd-El-Khalick 2012, 1055) and using prompts to get students to reflect on different issues related to the NOS (Akerson et al. 2000). Teachers can encourage students to discuss and exhibit their NOS understandings by planning for it, or through interactions during instruction such as questioning and conversations (Duschl and Gitomer 1997; Ruiz-Primo 2011; Shavelson et al. 2008).

3.3 | Providing Feedback That Moves Learners Forward

The effectiveness of formative assessment depends not only on the quality of elicited and gathered information about students' NOS learning but also on the teacher's subsequent actions (Bell and Cowie 2001; Black and Wiliam 2009; National Research Council 2001). Considering that a core component of NOS learning is to develop connections between different NOS ideas (Peters-Burton et al. 2019) and between NOS ideas and the context from which they are drawn (e.g., historical examples, socioscientific issues, and hands-on inquiry tasks) (Allchin et al. 2014; Khishfe and Abd-El-Khalick 2002), it is reasonable to assume that appropriate feedback and instructional correctives can play an important role in facilitating students' reflective learning of the NOS.

3.4 | Activating Students as Learning Resources for One Another

Formative assessment is a collaborative enterprise between teachers and students (Heritage 2013; National Research Council 2001). Although peer learning has received limited attention in the NOS literature, it can be an effective source of NOS learning through formative assessment. While peer

TABLE 1 | Overlapping objectives between formative assessment and explicit-reflective instruction.

	Where the learner is going	Where the learner is now	How to get the learner there
Teacher	*Clarifying, sharing, and understanding learning intentions* • Beginning a lesson with a statement of NOS learning goals (Khishfe and Abd-El-Khalick 2002) • Emphasizing why it is important learn about the NOS, particularly if the NOS do not feature strongly within the curriculum (Ryder and Leach 2008)	*Eliciting evidence of learning* • Helping students articulate their NOS understandings (Lederman and Abd-El-Khalick 1998) • Bringing students' tacit, fragmented and inarticulate NOS understandings to the forefront, and examine and revise them (Abd-El-Khalick 2012) • Using prompts to get students to reflect on different issues related to the NOS (Akerson et al. 2000)	*Providing feedback that moves learners forward* • Facilitating connections between NOS ideas and the context (historical episodes, socioscientific issues, and inquiry activities) (Khishfe and Abd-El-Khalick 2002) • Using a range of prompts and referring to several contexts (Abd-El-Khalick 2012; Allchin et al. 2014)
Peer		*Activating students as learning resources for one another* • Engaging students in evaluating peers' work and analyzing their own work in comparison to their peers' • Developing capacity to provide constructive feedback and internalize one's NOS learning	
Learner		*Activating students as owners of their own learning* • Providing structured opportunities to encourage learners to reflect on their science learning experiences in relation to NOS (Abd-El-Khalick 2012) • Supporting self-regulation and self-monitoring of students' NOS learning (Peters-Burton and Burton 2020)	

assessment is often discussed in the context of summative assessment, it can be facilitated in a more formative manner. Like the teacher, once the learning goal has been clearly shared, students can ask questions to each other (eliciting evidence of learning), evaluate and use the evidence to check for their peers' understanding as well as their own (see Section 5.3 for an example). Such student–student conversations, seen as informal formative assessment, can provide useful information about learning to the teacher. Research suggests that when learners engage in evaluating the work of their peers, they not only gain insights into alternative perspectives and approaches but also critically analyze their own work in comparison to that of their peers (Panadero 2016; Tseng and Tsai 2007). This reflective aspect of peer assessment can provide a powerful tool to encourage students to articulate and internalize their NOS learning experiences. In addition, as students actively participate in assessing and being assessed by their peers, they can develop a capacity for constructive feedback and refine their ability to identify gaps in their NOS understanding.

3.5 | Activating Students as Owners of Their Own Learning

Another important source of evidence is the learner's own assessment of their NOS learning, which can be supported by teachers providing "structured opportunities designed to encourage learners to examine their science learning experiences from within a NOS framework" (Abd-El-Khalick 2012, 1057). Given the centrality of reflection and self-regulation in NOS learning (Peters-Burton and Burton 2020), the use of formative assessment for NOS learning can have the potential to support learners in exerting personal agency and control over the learning goals. Importantly, Sadler (1989) argued that self-assessment is not something that happens naturally but is a skill to be learned. Teachers can support students' self-assessment by providing a protocol or log to record their performance, questionnaires, rubrics or checklists (Panadero et al. 2017), or more informally through interactions during the instruction. It is also important to note that self-assessment can be prompted by peer assessment and feedback (Heritage 2013). For example, if a student can share his or her idea about a certain aspect of NOS, a peer can elaborate or challenge it, and through discussion both can arrive at more sophisticated and justified ideas about the NOS. In this sense, the teacher can facilitate these peer and self-assessments as a form of reflective approach to NOS instruction.

3.6 | Summary

Although classroom assessment is an intrinsic element of explicit-reflective NOS teaching, how science teachers can collect and use evidence of students' NOS understanding during instruction to enhance learning has received limited attention. Drawing on the literature on formative assessment and NOS instruction, this section has articulated how formative assessment can support explicit-reflexive NOS teaching by focusing on how teachers elicit, interpret, and act on evidence of students' thinking in the flow of instruction. These activities can create opportunities for students' NOS views that are often fluid, tacit and fragmented (Leach et al. 2000) to be "brought to the forefront, examined and even revised through structured reflection," through discursive practices (Abd-El-Khalick 2012, 1055).

4 | Methods

4.1 | Curriculum Context

In the 2015 National Curriculum of Korea, a new subject 'Scientific Inquiry and Experimentation' was introduced as a stand-alone compulsory subject in 10th grade, to be taught alongside General Science. Its introduction was to help students "recognize the value of science as well as its impact on scientific inquiry and social, technological development" by learning about the NOS and scientific inquiry (Ministry of Education 2015, 25). This curriculum context provided a useful context to examine how teachers plan and deliver NOS lessons.

We focused on one school located in an urban and high-SES region in a metropolitan area in South Korea. The broader project focused on science teachers' NOS classroom assessment (including summative assessment) practices in this school in a time of curriculum reform. Among the data collected across two academic years from three teachers in the school, this paper focuses on Dr. Kim's (pseudonym) NOS lessons. The first author had known Dr. Kim for 3 years before the study, and from an informal conversation about how she was planning to teach and assess the new subject (described below), it was clear that her lessons would be a suitable site for a case study to investigate formative assessment that occurs in a NOS classroom. Unlike the other two teachers who described themselves as inexperienced with the NOS or its teaching, Dr. Kim had previous exposure to and a strong interest in NOS as a curricular construct. Given that NOS formative assessment is an underexplored and complex phenomenon, we decided to focus on Dr. Kim to provide illustrative and focused insights rather than generalizations, consistent with the purpose of case study research (Yin 2017).

Dr. Kim was an experienced earth science teacher with over 15 years of secondary science teaching experience and a PhD in science education, and had a long-term interest in the use of history and philosophy of science in her science instruction. This was in part because she found it difficult to do experiments in earth science which, by nature, deals with either historical or large-scale objects and phenomena, most of which cannot be brought to or replicated in classrooms. For her, using the history and epistemology of science was her personal strategy to incorporate scientific methods, inquiry and student-oriented learning. With such a background, she had knowledge of the NOS instruction such as the key aspects of the NOS, instruments, and the explicit-reflective approach. This is why we considered her an experienced teacher in the context of this study. In her assessments, she preferred using open-ended, free-response assessment items to multiple-choice items, because "kids who don't understand rarely write anything" (Interview 1). She found it problematic that the traditional image of science as being special and absolute had influenced the education of scientists as well as science teachers for a long time. However, in her view, having such a view of science is not helpful in today's world where "information

TABLE 2 | Summary of observed lessons.

Lesson	Outline
Lesson 1 (50 min)	• The teacher introduced the concepts of paradigm and theory change in science, using historical examples to illustrate these ideas. • The transition from Aristotelian to Galilean theories of motion was presented as a case study of paradigm shifts in science. • Students engaged in inquiry activities to explore how Galileo identified a paradox in Aristotle's theory of motion through a thought experiment on falling objects. • The teacher introduced key aspects of the Nature of Science (NOS), particularly how scientific knowledge evolves over time. • The group assignment was explained.
Between the lessons	• Students worked asynchronously for 2 weeks to prepare their group presentations before the next lesson.
Lesson 2 (50 min) Lesson 3 (50 min)	• Students were divided into groups of four and tasked with identifying an example of a paradigm shift in history. Each group prepared a 10-min PowerPoint presentation. All group members were asked to contribute to the presentation. • Following each presentation, the presenters answered questions from both the teacher and their classmates. • These two sessions were the analytical focus of this study.

is everywhere" and "proper decision-making requires the capacity to assess a given piece of information from various angles and with critical mind" (Interview 3). She said that such an ability is difficult to cultivate while teaching the content knowledge-focused curriculum, so it should be fostered through teaching the NOS.

Among the six NOS-related lessons that Dr. Kim taught throughout the semester, we selected three sessions focused on the development of science for a close examination. These lessons involved whole-class conversations about the NOS, including teacher–student and student–student interactions, which provided an ideal context for examining informal formative assessment of the NOS and how the teacher facilitates it. The other three NOS lessons either included an online component (due to the partial transition to online instruction) or were not designed to facilitate much interaction (also due to the COVID guidelines), making them less suitable for conversation analysis. Table 2 summarizes the three lessons that were observed.

At the beginning of the semester, Dr. Kim taught three 50-min lessons on the development of scientific knowledge. This was stated in the national curriculum as an explicit learning goal ("Understand scientific discoveries that motivated paradigm shifts in the history of science and explain how science develops"), although it did not provide specific pedagogical guidelines. In the first session, she started by explaining that the term, paradigm, refers to a system of thinking that defines people's opinions and thoughts in a certain era. Then she highlighted that while science sometimes develops incrementally and continuously, at other times the development occurs drastically, accompanied by a shift of paradigm. She elaborated this key NOS aspect by referring to Galileo's thought experiment on motion that triggered the shift from the old Aristotelian paradigm to the new era of classical mechanics. After this introductory session were two sessions devoted to a project-based assessment task, which were the analytical focus of this study. Students were divided into groups of four to find a historical example of a paradigm shift and prepare a 10-min PowerPoint presentation, and worked in groups asynchronously for 2 weeks. After each group's presentation, Dr. Kim initiated a whole-class discussion on the presentation, and facilitated and orchestrated the informal assessment conversation. More details about her pedagogical reasoning involved in the lesson and assessment design will be explained in Section 5.1.

Furtak et al. (2019) conceptualize formative assessment as "a network of *participants* engaging in *practices* organized and coordinated by a set of *tools* within a specific context" (p. 99). In light of this conceptualization, this study's context includes the *tools* for formative assessment (i.e., the task of preparing and delivering a presentation in groups), the *practices* for it (e.g., teacher's and students' uses of talk moves to probe each other's thinking, provide feedback, and press for reasoning), and the *participants* (i.e., the teacher and students as sources of NOS learning). We viewed Dr. Kim's project-based NOS task as a form of *formal* formative assessment. During class discussions, however, she also responded spontaneously to students' ideas and questions, using them to check understanding and create learning opportunities—an example of *informal* formative assessment (Duschl and Gitomer 1997). Her lesson thus incorporated both forms, each supporting NOS learning by eliciting, interpreting, and acting on evidence of student understanding (Ruiz-Primo and Furtak 2007).

4.2 | Study Design and Data Collection

This study utilized a naturalistic qualitative case study approach, with the aim of closely examining the complexities arising from one experienced teacher's classroom practice. We see this study as a revelatory case study given that teacher practices in NOS formative assessment are a largely underexplored phenomenon (Yin 2017). Our larger project aimed to investigate science teachers' practices of classroom assessment in the wake of the national curriculum reform in South Korea that highlighted NOS as an explicit learning goal. Ethics approval for the study was granted by the Research Ethics Committee at the Department of Education, University of Oxford.

Data sources included observations of Dr. Kim's three lessons (50 min each, video recordings and field notes), related artifacts (e.g., textbook, Dr. Kim's lecture slides, and students' presentation slides), and interviews with Dr. Kim. For the lesson observation, while a standardized observation rubric was not used, the observation focus was informed and guided by the five strategies of formative assessment described in the literature review. Observational attention was given to how learning goals were communicated, how evidence of NOS learning was elicited, the types of feedback provided, and moments of peer and self-assessment (Leahy et al. 2005). The observation notes guided follow-up interviews with her, as well as generated initial analytical ideas that were revisited later during the analysis. The national curriculum and the school's instruction and assessment plans for the subject were also collected as documentary sources.

The first author conducted three semi-structured interviews with Dr. Kim after each of the three lessons. The first interview focused on her general views about science teaching and assessment as well as her knowledge and experience related to the NOS specifically (e.g., How would you describe your approach to science assessment? How would you define NOS? Do you teach NOS in your science class?). In the second and third interviews, Dr. Kim was asked to provide her own reflection about the lesson (e.g., How did you feel about the lesson?) and was also asked about specific elements of her pedagogical actions, intentions, and reasoning that the first author noticed (e.g., You mentioned that 'This class is not about the scientific knowledge'. Can you explain what you mean by that?). Each interview lasted between 60 and 90 min. Throughout the research, there were a number of other informal communications with Dr. Kim in the form of casual conversation over coffee (recorded in the field notes) and extended text messages, some of which were included in the analysis with her consent. Interviews were used to triangulate and contextualize the classroom observations. While the lesson recordings captured in-the-moment formative assessment interactions, the interviews allowed us to explore Dr. Kim's instructional intentions, pedagogical reasoning, and reflections on the effectiveness of her strategies, and her interpretation of specific classroom events. This triangulation helped to ensure that observed practices were not interpreted in isolation, but situated within the teacher's broader epistemic and pedagogical approach.

4.3 | Data Analysis

Lesson recordings (Lessons 2 and 3), interviews, and related artifacts and documents were transcribed verbatim and imported into MAXQDA (VERBI Software 2019) for qualitative analysis conducted over three rounds. Although we broadly followed three coding cycles, the analysis process was more iterative than linear. Initial interviews and informal conversations held prior to lesson observations were analyzed first, in order to understand the teacher's background and the curricular and instructional context outlined in Section 4.1.

Given that all observed lessons followed a single pedagogical structure (student group presentations followed by whole-class discussion), this study is best understood as a case of an experienced teacher's formative assessment practices within a specific, project-based instructional design. This narrow focus allows for a deep analysis of the in-the-moment moves, opportunities, and challenges associated with this particular approach to teaching the NOS.

4.3.1 | Round 1: Deductive Coding

For the first cycle of coding, we used Ruiz-Primo and Furtak's (2007) ESRU model to identify the structure of the informal formative assessment occurring in the teacher–student, student–student interactions during the whole-class discussions. The ESRU model was proposed as a way to capture differences in teachers' informal formative assessment practices (Ruiz-Primo and Furtak 2007). It has been suggested that the traditional models such as the IRF model (initiation, response, and follow-up; Sinclair and Coulthard 1975) and the IRE model (initiation, response and evaluation; Mehan 1982) are more suitable for lessons where teacher talk is predominant and the "correct response" is pursued (Ruiz-Primo 2011) than for lessons with frequent participation of students and open-ended lesson topics such as NOS. These characteristics of the ESRU model made it a suitable framework to analyze Dr. Kim's lessons, which involved as much student talk as teacher talk, and where the lesson goal was to understand the NOS.

Table 3 shows a short example of ESRU coding. This excerpt followed Group 1's presentation on the discovery of the double helix structure of DNA as an example of a paradigm shift. She starts by asking a question to *elicit* the presenters' ideas about why their example (i.e., the double helix example) was paradigm-shifting (Line 2). A 10-s silence followed, which was coded as a *student response* (Line 3). Coding silence as a form of response was considered reasonable given that it still provides the teacher with some information about what students know or do not know. Hyun then gives *another response* to Dr. Kim's question based on the utility of the discovery (Line 4). Dr. Kim then briefly *recognizes* his response with "Yes" (Line 5), and *uses* the group's insufficient response to reiterate what a paradigm is and provide a better answer to her own question (Line 6). When the group agreed to her interpretation, she *recognizes* their selection of the example and moves on to initiate another elicitation.

This first cycle coding enabled preliminary sorting of assessment conversations, and importantly, the identification of "episodes." Since the assessment conversations occurred after each group's presentation, an episode usually started when a peer or the teacher asked the presenters to clarify, explain or elaborate on what they said, and ended with the teacher using the information to take actions. A total of eight formative assessment episodes were identified in the transcripts, each consisting of complete or incomplete cycles. This deductive coding served as a preliminary step to discovering discursive meanings, rather than quantifying the cycles or their patterns as was the case in several other studies (Nieminen et al. 2016; Rached and Grangeat 2021; Ruiz-Primo and Furtak 2007). For example, we noticed that a certain category of teacher questions repeatedly

TABLE 3 | Round 1 and Round 2 coding example.

	Speaker	Utterance	Round 1 coding (deductive)	Round 2 coding (inductive)
1	Dr. Kim	... So today's presentation was meant to be about finding and introducing a historical inquiry that led to a paradigm shift.		States learning goal
2		So I'll ask a question first. It applies to all groups. Why do you think this is a historical case that shows paradigm shift? I'd like anyone from the group to answer.	Eliciting	Helps to relate the example with NOS
3		(Silence, 10 s)	Student response	
4	Hyun	When we know the helical structure of DNA, it makes opportunities for human cloning, and tomato bearing potatoes, and the like, so we can say it's a paradigm.	Student response	
5	Dr. Kim	Yes.	Recognizing	

failed to prompt students' reactions or initiate an assessment conversation cycle (see Section 5.4), which signaled the need for further investigation. This allowed us to examine not only how often formative assessment occurred, but also what kinds of discursive contingencies made certain episodes pedagogically productive or unproductive for NOS learning.

4.3.2 | Round 2: Inductive Coding

Although the ESRU cycle is useful in characterizing the discursive structure of whole-class conversations and the role of each statement within that structure, it can provide limited information as to how the statement relates to the teacher's NOS instructional intentions and strategies. In the second-cycle coding, conversations in each episode were coded inductively, with a focus on how the actions of the teacher and the students help or hinder explicit and reflective instruction of the NOS (Table 1). For each utterance, analytic questions such as "Who is saying what and how?" (Tang 2021), "What is the intention of this talk?", "How does it connect to what comes before and after?", and "What purpose does this talk serve in terms of NOS teaching and learning?" were asked. When the context and intention of the talks were not evident from the text, we returned to the recording and watched it again to gain a more nuanced understanding. Analytic memos and the logbook function in MAXQDA assisted with keeping track of how the analysis evolved and how each coding and categorization decision was justified (Birks et al. 2008).

An illustrative example of inductive coding is provided in the rightmost column of Table 2. As compared to deductive coding, the codes here are more representative of the purposes of the talk made by the teacher, informed by the five formative assessment strategies discussed in Section 3 (Leahy et al. 2005). Since the lesson started directly from the presentation without introductory remarks by the teacher, she starts by *clarifying the learning goal*, which was to learn about the historical examples that show a paradigm shift. Her next talk included *explaining the NOS concept* of interest (i.e., what a paradigm is) and *helping students connect their example to the target NOS aspect*. Then she *made a compliment* about their choice of the example. As such, the second cycle allowed contextualizing each utterance within NOS learning and characterizing its purpose and functions, adding nuances to the ESRU analysis. Note that the students' (both presenters and the audience) role was not limited to merely responding to the teacher's questions; they asked questions, recognized others' work, and helped each other during the whole-class discussions, which contributed to communal learning. In this way, the inductive analysis was used to qualify, extend, or in some cases challenge the implications of the ESRU cycles. For example, a cycle might be structurally "complete" in ESRU terms, but still fail to advance NOS learning if the teacher's feedback lacked reflective focus—something only visible through the inductive lens.

4.3.3 | Round 3: Developing Themes

Once all episodes were coded in two rounds, we looked for patterns across episodes using the constant comparison method (Glaser and Strauss 2017). In this round, we examined interviews and other data sources alongside the Round 1 and Round 2 coding results to develop an aggregated interpretation across the dataset. In doing so, the ESRU structure provided a scaffolding to track the interactional dynamics of formative assessment, while the inductive coding revealed how particular moves enabled or constrained NOS learning. The interview data was read through and coded in parallel with the lesson transcripts, with the aim of understanding Dr. Kim's views, reasoning, and intentions.

The initial inductive codes from Round 2 were adjusted, relabeled, and assembled into four categories that characterized Dr. Kim's instructional strategies. These strategies were not imposed a priori but emerged from repeated attention to what pedagogical functions the teacher's and students' moves served within NOS instruction. The deductive coding from Round 1 served two purposes in this process: (a) it allowed the instructional strategies to be interpreted as formative assessment strategies by paying attention to how evidence of learning is elicited and acted upon in assessment conversations (Duschl and Gitomer 1997); (b) with the cyclical focus of the ESRU method, we were able to identify the discursive patterns (e.g., whether an elicitation was followed by subsequent phases, who initiated and who terminated a cycle, in what order evidence was elicited and acted upon). For example, in the assessment conversation presented in Table 9, the deductive coding indicated that the cycles were largely incomplete, suggesting challenges with the particular strategy used by Dr. Kim.

4.3.4 | Trustworthiness

The coding took place through a collaborative and iterative process. The first step was to familiarize the analyst with the data (Saldaña 2013; Silverman 2014). After importing the data into MAXQDA, the first author undertook the initial coding of the data as described above. The coded data and emergent themes were subsequently reviewed by the second and third authors in monthly meetings. Their role was to critically evaluate the first author's coding decisions, challenge interpretations where necessary, and provide alternative perspectives. We discussed what each of us saw in the transcripts, what emerges from our different perspectives and any alternative interpretations of the data (Morse 1997), discussing any discrepancies to agreement. Reliability was established through these iterative cycles of peer scrutiny and collaborative sense-making, aiming for "dialogical intersubjectivity," that is, "agreement through a rational discourse and reciprocal criticism between those interpreting a phenomenon" (Brinkmann and Kvale 2015). Coded excerpts, memos, and emerging themes were shared and discussed regularly with the second and third authors, who challenged interpretations, raised alternative readings, and helped refine the analytical logic. Where there were uncertainties or divergent interpretations (e.g., how a particular utterance functioned as feedback, elicitation, or both), these were resolved through reexamination of the raw data and collaborative discussion until consensus was reached. We triangulated our data analyses from different sources (interviews, informal communication, and lesson observations) to enhance validity. A codebook with inductive codes and their definitions was generated (Table 4) to support transparency and shared understanding within the research team (Oliveira 2023).

5 | Results

As a result of the three rounds of coding, we identified four distinct strategies for formative assessment that Dr. Kim enacted to facilitate explicit and reflective NOS learning: *multicontextualizing NOS, mapping examples to NOS, focusing discussion on NOS*, and *exploring different aspects of NOS*. These strategies were observed in the six whole-class discussion sessions, sometimes initiated by Dr. Kim's pedagogical moves (e.g., asking questions), planned or spontaneous, but at other times by the questions and comments from peers, voluntarily or after being called on by Dr. Kim, showing the diversity of formative assessment that happened in her classroom. In what follows, to address the two research questions (1. How did the teacher's formative assessment strategies create opportunities for NOS learning? 2. What challenges did the teacher encounter in using formative assessment to support NOS learning?), we use episodes to illustrate how these strategies manifested in her interactions with the presenting group as well as the whole class, what kind of moves contributed to each of her strategies, and also what complexities and challenges of the teaching and formative assessment of the NOS emerge from the episodes.

5.1 | Multicontexualizing NOS: Examining the Paradigm Shift in Different Contexts

Dr. Kim's first strategy was embedded in her planning of the assessment task, in which she utilized multiple contexts to shed light on a single aspect of NOS. The primary reason for using this group task, she said, was to exemplify NOS in a range of historical contexts (Interview 2). She believed that this task would also increase student participation as they work in groups and share the results with each other. In the post-lesson interview, she clarified her reasoning for having multiple groups present on different topics:

> After I taught NOS explicitly with the PowerPoint slides I had made, [I wanted to address] historical examples that bear those NOS aspects. *It'd be nice to talk about one [example] in depth*, but then it could have been only about me [teaching it to students]. The new subject has limitations in transmitting [scientific content] knowledge. *So when they work together, they could encounter various topics of inquiry, various historical examples, so I used the activity.* (Interview 2; all emphases were added by the authors)

This quote suggests that Dr. Kim used the task to make students' NOS understanding visible. It also points to the choice that she had to make between breadth and depth in designing a NOS assessment task. She acknowledged that either approach has values, but she chose to prioritize breadth for two reasons: to increase student participation and to allow for some student-led learning of content knowledge of science alongside NOS. In other words, she multicontextualized NOS by setting out to examine how a certain aspect of NOS can be observed in a range of different contexts. After the final group's presentation, Dr. Kim opened the discussion by reminding students of her instructional intention:

> So, after all, you had an opportunity to *think about the [content] knowledge and the nature of science from at least six examples*. I hope you have learned something from them.

Indeed, students presented a range of historical scientific discoveries, although not all were clearly examples of a paradigm shift. Some groups selected cases that more closely aligned with the concept of a paradigm shift, while others chose examples that were less clearly connected. However, Dr. Kim did not explicitly draw attention to these contrasts. She responded to each group's presentation individually, offering comments or brief affirmations, but did not initiate a class-wide discussion that linked the

TABLE 4 | List of inductive codes from Dr. Kim's paradigm shift lesson.

Dr. Kim's instructional strategies			
Theme	**Code**	**Description**	**Example**
Multicontextualing NOS	Uses multiple historical episodes to teach paradigm shift	Teacher designs and enacts an assessment activity where students present diverse historical episodes to illustrate paradigm shifts.	Design and use of the assessment task where students learn various examples of paradigm shifts from history of science.
	Clarifies learning goal	Teacher states the goal of understanding paradigm shifts from multiple contexts.	Teacher: Good. Well done. … In the end, you had an opportunity for thinking about the (content) knowledge and nature of science about at least six examples. I hope you could learn something from them.
Mapping examples onto NOS	Clarifies learning goal	Teacher reminds students of the goal: understanding historical cases that led to paradigm shifts.	Teacher: So today's presentation was meant to be about finding and introducing a historical inquiry that led to a paradigm shift.
	Explains a NOS concept	Teacher explains what a paradigm is.	Teacher: Yes. A paradigm is, easily speaking, a system of thinking that prevailed people's thoughts in a certain era.
	Helps relate the example to the paradigm shift	Teacher asks or explains how the example demonstrates a paradigm shift.	Teacher: So I'll ask a question first. It applies to all groups. Why do you think this is a historical case that shows a paradigm shift? I'd like anyone from the group to answer.
	Gives a compliment	Teacher asks or explains how the example demonstrates a paradigm shift.	Teacher: So well done on presenting the continental drift. Then we have Q&A. I can't give you much time, so just 1 min, let's get one question. You explained really well about the paradigm shift, and connected it with continental drift so I want to clap.
Focusing the discussion on NOS	Delivers relevant content knowledge	Teacher provides scientific knowledge necessary for understanding NOS.	Teacher: So Einstein's strong theory declines, and cosmology is, it is what uncovers the origin and evolution of the universe. In so far as the origin, and the evolution process, people have made remarkable development since. So we know the Big Bang theory is reliable among cosmological theories, and how the universe expands, with acceleration or otherwise, people are doing research about that. There are some conclusions out there. In that regard, it is clearly an example of a paradigm shift, right?
	Clarifies learning goal	Teacher articulates the NOS focus of the lesson.	Teacher: So now, we are not focusing on the scientific knowledge to learn about Kepler's third law or specifics of Newton's law, so we can move on. But what's important is, when you present, it's important that you summarize and present what you can digest, right? In that regard, although it's not a knowledge-focused presentation, you need to be able to introduce it [relevant scientific knowledge] briefly. The part you are talking about.

(Continues)

TABLE 4 | (Continued)

Dr. Kim's instructional strategies			
Theme	**Code**	**Description**	**Example**
	States expectations about content knowledge	Teacher indicates expectations for how much scientific content should be presented.	Teacher: … It would've been nice if you focused on that and presented a bit easier. All groups alike. Let me tell you. Only the things you guys can understand. You don't need to bring all the evidence. Only one of the three, but it's good if you digest it and express it in your own language. Anyway I know that you prepared a lot …
	Shuts down detailed content questions	Teacher discourages or shuts down questions on scientific details.	Student: … Could you explain what Kepler's planets are about? Teacher: So now, we are not focusing on the scientific knowledge to learn about Kepler's third law or specifics of Newton's law, so we can move on.
Exploring diverse NOS aspects	Explains a NOS concept	Teacher introduces additional NOS aspects for consideration.	Teacher: Then in relation to tentativeness, the nature of science is that, science is technology-dependent. Or one development in science and technology can bring about another development. You can say these things.
	Helps relating the example to other NOS aspects	Teacher scaffolds student understanding by linking to other NOS aspects.	Teacher: Any questions related to that? (Silence for a sec) Is it too difficult? Do you need another summary? When we studied with the electricity example, don't you recall anything? Something leads to something. The blank filling that you like. Hint, the development of science … (Students silent) (Laughter). So the limitation of plate tectonics, will be able to be overcome with a new theory, through the development of the exploration techniques. Right? (Yes.)
	Connects to previous learning	Teacher references prior learning on science content, NOS, or historical examples.	Teacher: When we studied with the electricity example, don't you recall anything? Something leads to something. The blank filling that you like. Hint, the development of science …

six examples or guided students to make cross-case comparisons in relation to the intended learning goal. As such, the opportunity to synthesize ideas across the different contexts and reinforce the NOS concept of paradigms remained underdeveloped.

5.2 | Mapping Historical Examples Onto NOS: More Than "What Happened in the Past"

The explicit-reflective approach to NOS asserts that learning about stories from the history of science can be useful but is not sufficient for students' NOS learning (Abd-El-Khalick et al. 1998; Akerson et al. 2000). In other words, NOS is not naturally learned by simply learning about what people discovered in the past and how. When asked about the purpose of assessing NOS learning during the initial interview, Dr. Kim emphasized the importance of students being able to connect abstract NOS ideas to concrete cases, such as scientific inquiries or historical episodes:

> [The purpose of assessment is] how well one understands the NOS. That's it … Not just saying it blah blah only in words, but being able to *apply what they understand when doing something like inquiry examples or historical examples*. "This [case] is indeed about the tentative NOS, this part relates to what," if they can say things like these. (Interview 2)

Here, Dr. Kim was articulating her instructional goal that students should not merely recite definitions of NOS concepts but demonstrate their understanding by identifying how these ideas—such as the tentativeness of scientific knowledge—are demonstrated by specific examples. This reflects her commitment to an applied, contextualized approach to NOS teaching and assessment.

Throughout the lessons, we identified several of Dr. Kim's pedagogical moves to keep the students reminded of the lesson goal, which was to understand the historical discoveries that led to a shift of a paradigm. The discussion that occurred immediately after the first group's presentation on the discovery of the double helix structure of DNA illustrates this effort. In their presentation, Group 1 used cartoon clips that featured four major figures in the story, James Watson, Maurice Wilkins, Francis Crick, and Rosalind Franklin, to illustrate the story of the double helix structure. They then ended by drawing attention to the neglected contributions of the female scientist Franklin in the discovery process. In the following transcript, Dr. Kim opened the whole-class discussion by asking a question to the presenters (Table 5).

In this initial set of interactions, Dr. Kim's intention to focus the discussion on the "paradigm shift" and the development of scientific knowledge (i.e., the NOS lesson objective) is made clear. First, she tells the class what the given task was. Then she asks a question that encapsulates such an intention: "Why do you think this is a historical case that led to a paradigm shift?," which she had planned in advance to ask all groups (Line 2). It took the group, standing in front of the class, about 10 s until one of them started answering this question. During that pause, they turn their heads toward each other, shrug their shoulders, one by one, as if they are waiting until any one of them will say anything, exchanging mixed facial expressions (Line 3).

TABLE 5 | Excerpt illustrating how Dr. Kim mapped historical examples onto NOS (1).inte.

	Speaker	Utterance
1	Dr. Kim	... So today's presentation was meant to be about finding and introducing a historical inquiry that led to a paradigm shift.
2		So I'll ask a question first. It applies to all groups. *Why do you think this is a historical example that led to a paradigm shift?* I'd like anyone from the group to answer.
3		(Silence, 10 s) (The presenting group stares at each other and shrugs)
4	Hyun	When we know the helical structure of DNA, it makes opportunities for human cloning, and tomato bearing potatoes, and the like, so we can say it's a paradigm.
5	Dr. Kim	Yes.
6		A paradigm is, easily speaking, a system of thinking that prevailed people's thoughts in a certain era.
7		Then you mean that there was a huge change in biology before and after discovering the double helix structure?
8	Lin	Yes.

From these 10 seconds, Dr. Kim notices that the group is not prepared to answer her question and has not thought about it before. After an awkward pause, Hyun says that it was a paradigm shift due to the practical benefits of the advances in genetics to human lives (Line 4), which is not a defining characteristic of a paradigm that Dr. Kim had taught about in the previous session. Nevertheless, instead of criticizing or correcting their answer, she quickly recognizes them by saying "Yes." She then goes back to reminding the class of what a paradigm is, which is a prerequisite to connecting the example to the paradigm shift. She provides, in the form of a question, how the presented case about DNA can be interpreted as a paradigm shift. Although this interaction looks as if she is revoicing students' reasoning, particularly with her use of "you mean ..." (Line 6), it appears that the group had not reasoned that way until Dr. Kim did it for them (Line 7). The group then agrees with Dr. Kim's reasoning (Line 8), and she gives a compliment by acknowledging their selection of the example (Line 9).

Although the above episode seemed to have resolved the issue of linking the DNA story to the paradigm shift idea, it soon turned out that the debate was not over. This became clear when June, from the audience, immediately after Dr. Kim's feedback, asked a question:

June: Since it's about the paradigm shift, then what's the paradigm that existed before the double helix?

(Class laughs)

Hyun: I don't get the question.

June: So, when the DNA's double helix structure was discovered, your topic is that it is a paradigm shift. Then what was the paradigm before it?

Whether the DNA example represents a paradigm shift is not straightforward—it can be seen either as a discovery within an existing paradigm or as marking molecular biology's emergence as a new one. Interestingly, Dr. Kim did not address this in class, though she allowed the discussion to continue. Later, in conversation with the first author, she acknowledged that both the textbook and her lesson described "paradigm" too broadly and vaguely. Following the national curriculum, she defined it as "a system of thinking that prescribes the opinions and thoughts of people living in a certain time" (PowerPoint slide) and treated it as a mechanism through which science advances. Yet neither source clarified what actually counts as a paradigm shift—a common simplification that occurs when specialist knowledge is didactically transposed (Kampourakis 2016).

Similarly, the whole-class discussion after Group 4's presentation of the Mpemba effect provides an example of peer formative assessment against the shared learning goal, but in a more sophisticated form. The Mpemba effect refers to a phenomenon that hot water takes less time to freeze than cold water does, coined after a Tanzanian schoolboy Erasto Bartholomeo Mpemba who discovered it while making ice cream (Mpemba and Osborne 1969). Group 4 also spent most (8 m 33 s of 11 m

57 s) of their presentation time to give a detailed scientific account of the Mpemba effect itself. The group started by explaining how the effect came to be known and has been verified and explained by scientists until recently. To do this, they explained the molecular structure of water, and how water molecules are tied together with hydrogen bonds.[1] They then explained that hot water molecules can release the energy stored in hydrogen bonds faster than cold water molecules do, hence freezing faster, and concluded the presentation.

Dr. Kim asked the presenters to moderate the whole-class discussion, and they said if there were any questions, but no question was asked. Dr. Kim then looked at the first author sitting at the back of the classroom observing the discussion, and asked if he had any. He recognized the group's interesting presentation and then asked for clarification on why the Mpemba effect example makes a paradigm-shifting discovery. The answer they gave the researcher then faced a series of questions from peers. This process is illustrated in the following episode (Table 6).

Hyuk's response to the researcher's question, which would not have been a surprise by this moment, begins by referring to the definition of a paradigm. He then uses this criterion to claim that an old, common idea that existed since Aristotle's time was changed by the 2003 discovery that explained the Mpemba effect, so it counts as a paradigm shift (Line 3). Sun then asks a question to check if they said that correctly about the average person's idea (Line 4), and Hyuk confirms that it is what he had meant (Line 5). Sun then repeats the question (Line 7), and now Hyuk tries to more precisely distinguish the old state and the new state in his response based on whether the phenomenon was "proved" (Line 8). Despite their answer, Jin is still unsure and asks if the effect was indeed already known from a long time ago (Line 9). Sun adds to her comment using a conditional statement, saying that the Mpemba effect must have been previously unknown to count as a paradigm-shifting discovery (Line 10). Jin then uses Sun's criterion to assess Hyun's answer—people knew the phenomenon was already known, so it was not a paradigm-shifting discovery (Line 11).

As it becomes gradually clear that the Mpemba effect itself is not a paradigm shift example, Beom picks up on the word "prove" from Hyuk's answer and asks if it was why they selected this example (Line 12). Hyuk states that the addition of proof indeed was the paradigm-shifting part since it is what they thought to be the most important contribution (Line 13). Although the group's intention became less confusing by changing the focus from the phenomenon itself to the recent scientific verification of it, Dr. Kim says that the Mpemba effect was not a good example of a paradigm shift (Line 15). Although she could have ended the discussion here and moved on to the next group, she made an attempt to use the group's presentation to support further learning, which we will return to later in Section 5.4.

These exchanges between Hyuk and the audience illustrate how peer assessment can facilitate students' sophisticated understanding of the NOS. Without any help from the teacher (although initiated by the researcher's question), the students in the audience recognized the problems of linking the Mpemba effect to the paradigm shift and brought this issue to the forefront, asking the presenting students to clarify their idea and examine it. In the meantime, the condition of "being previously unknown" emerged and was subsequently articulated as the whole-class discussion moved on. These can be seen as examples where not only the teacher but also learners become agents of formative assessment (Wiliam 2010), and through such engagement, reflect collaboratively on and construct knowledge about the NOS.

5.3 | Focusing the Discussion on the NOS: The Dilemma of Addressing Content Knowledge

Dr. Kim's effort to keep the lesson focused on the NOS also unfolded in another related context, as she was trying to strike a balance between a NOS aspect and the scientific content knowledge from which the NOS aspect is derived. This strategy was manifested in Dr. Kim's talk as well as how she managed the formative assessment conversations during whole-class discussions.

Group 3 presented on the shift of paradigm from the geocentric to the heliocentric view of the universe. The group spent about 1 min and a half explaining the backgrounds and contents of Ptolemy's and Copernicus's models. A member of the group picked up a marker and started drawing the Sun, Venus, and the Earth on the whiteboard to explain how the phase and size change of Venus supports heliocentrism over geocentrism. They then went on to briefly (less than 30 s) introduce that there were three major pieces of evidence that supported the heliocentric model: Kepler's three laws, Galileo's observation reports of Jupiter, and later studies by Newton and other scientists. The presentation ended with summarizing the two systems and key aspects of the change, and sharing the group's reflection on their collaboration.

In the whole-class discussion that followed, Eun raised her hand and asked what she called a "specific" question about the presentation, as shown in the following excerpt (Table 7).

In this episode, Eun became curious about a particular aspect of the group's presentation, namely Kepler's laws, and asked if they could elaborate on it (Line 4). Before she asked this question, she checked if it would be okay to ask this kind of "specific" question to the group, signaling that it might be tricky and caused class laughter (Lines 1–2). Won, from the presenting group, takes the challenge by saying "Well, if I can." (Line 3). At this moment, we can see that this whole-class discussion has turned into a test of the presenter's ability to answer a hard question, as if it were a kind of game show. When Won, with no hesitation, started saying "First of all …" (Line 5), the class got excited and said "oooooh" (Line 6). However, to the disappointment of his peers, the response that he made was merely a repetition of his earlier argument that Kepler's laws were part of the evidence for the heliocentric theory, and did not really answer what the laws actually are. The class realizes this and reacts with laughter and a friendly boo, showing their disapproval (Line 8). The presenter admits that he did not answer the question and recognized the gap in his knowledge, by saying that the content was quite a difficult one (Line 9).

TABLE 6 | Excerpt illustrating how Dr. Kim mapped historical examples onto NOS (2).

	Speaker	Utterance
1	Dr. Kim	… (Asks the researcher) Do you have a question? We can have a critical question.
2	Researcher	[To presenters] I liked your presentation. I just missed what you said about *how this example relates to the paradigm shift, so could you explain that again?*
3	Hyuk	So, this is a paradigm. A paradigm means what people think over a time period, until now, for as long as 2000 years. An average person would think so [that hot water would freeze faster]. But like in the 2003 discovery [theoretical verification of the Mpemba effect], breaking the stereotype, so breaking it and changing people's thinking so we thought it's a paradigm shift.
4	Sun	*So in the past, people took it for granted that hot water freezes faster than cold water?*
5	Hyuk	Yes, didn't they?
6		(Students in the audience look confused and murmur to each other)
7	Sun	They knew that since long ago?
8	Hyuk	Long time ago, Descartes and Aristotle mentioned such a thing. But they couldn't *prove* it and just thought that it is what it is. But through the 2003 discovery …
9	Jin	Then isn't it that they already knew it? That it freezes faster?
10	Sun	So if they had not known it [hot water] freezes faster, it [the discovery of Mpemba effect] would have been a paradigm shift …
11	Jin	Well, they knew it. It's not like they came to learn about a new theory.
12	Beom	You mean they hadn't *proved* it?
13	Hyuk	Yes, because the proof was not there until …. *That's right. We think so. Anyone can make a hypothesis, and anyone can have thoughts. But we think that publishing [sic: proving] is important, so we presented with that focus about the paradigm shift.*
14		(Applause)
15	Dr. Kim	I think connecting this example directly to a paradigm shift is a bit far-fetched.

Dr. Kim's comment after this "game" ended has several important implications. With her opening sentence, she is stressing that the point of their discussion is not the specifics of the evidence or any theories, thereby refocusing the lesson on the target NOS (Line 11). Second, with "we can move on" she intends to end the discussion on this "specific" question about content knowledge without further elaboration. In the rest of her turn (Line 13), she tells the class that "scientific knowledge" still matters although it is not the focus. She then states her expectation pertaining to the presentation of content knowledge—that students are expected to summarize and present the relevant content knowledge that they can "digest" and mention in the presentation. She concludes this discussion by acknowledging the group's effort. Dr. Kim's decision to redirect the conversation toward the epistemic focus ("we are not focusing on the scientific knowledge…") demonstrates her responsive use of evidence gleaned from student talk. By interpreting the peer-elicited question as an indication of where students' attention had drifted, she adjusted her instructional strategy in real time—an example of formative decision-making to maintain coherence with the NOS learning objective.

This episode, catalyzed by a "specific" question from a peer, shows the difficulty of dealing with the relationship of content knowledge and NOS in science lessons. Questions about scientific details from peers happened again while discussing Group 5's presentation, but Dr. Kim immediately shut the conversation down ("Don't ask hard things like the critical value"), explaining relevant content knowledge on her own. Later in the interview, she recalled that she had "stopped them when they ask too difficult science-related questions" and "explained something that students can't explain themselves rather clearly and easily" (Interview 3).

5.4 | Exploring Diverse NOS Aspects: Thinking Beyond the Target NOS

Throughout the paradigm shift lesson, there were three instances in which Dr. Kim elicited her students to think of NOS aspects besides paradigm shift from the examples that the groups presented. However, none of her elicitation attempts resulted in students' responses, so she ended up providing the additional NOS aspects on her own. In terms of the ESRU cycle, no cycles were observed as there was no response from the students.

One such effort was observed after Group 4's Mpemba effect presentation. After the peer assessment conversations where the paradigm shift aspect of the story was debated (see Section 5.2), Dr. Kim shut down the discussion momentarily by telling the

TABLE 7 | Excerpt illustrating how Dr. Kim kept the discussion's focus on the NOS.

	Speaker	Utterance
1	Eun	Can I ask about a *specific* thing, ask specifically? I can't, right?
2		(Class laughs)
3	Won	Well, if I can [answer it].
4	Eun	Frankly, it's out of personal curiosity. So when I see your PPT, the evidence for heliocentric theory include Kepler's planets, Galileo's observations and the like. *Could you explain what Kepler's planets are about* (Note: The students had not learned about Kepler's laws of planetary motion in the National Curriculum.)
5	Won	First of all, Kepler ...
6	Audience	(With excitement) Ooooooh
7	Won	... argued for heliocentric, heliocentric theory. Then many scholars did various research to prove the heliocentric theory, and among them, there was Kepler's third law of planetary motion ...
8		(Class laughs and boos)
9	Won	The content is quite hard.
10		(Class laughs)
11	Dr. Kim	So now, *we are not focusing on the scientific knowledge* to learn about Kepler's third law or specifics of Newton's law, so we can move on.
12		*But what's important is, when you present, it's important that you summarize and present what you can digest, right?*
13		In that regard, although it's not a knowledge-focused presentation, *you need to be able to introduce it [relevant scientific knowledge] briefly*. The part that you are talking about. So good, well done.

class that it would be "a bit far-fetched" to call it a paradigm-shifting discovery. However, she did not end the discussion there. Instead, she attempted to use their example to discuss other NOS aspects than the paradigm shift:

Dr. Kim: I think connecting this example directly to the paradigm shift is a bit far-fetched. But it's still a meaningful example in terms of the nature of science. Bin, do you want to talk about it? What aspects of the nature of science...

(Silence)

Dr. Kim: Or Jung or Young? Bin has talked a lot.

(Silence)

Dr. Kim: So first, we can see, scientific knowledge is confirmed by evidence, through an example. Next, they developed a system for an explanation. Inquiry techniques had developed over time, so it [Mpemba effect] was proved, right? As we know more, the more things we can prove, and science develops. Science is based on, relies on existing theories. We can say these things.

Here, she starts by recognizing the value of the example for learning NOS ("... it's still a meaningful example"). Then she calls on a student in the audience to ask what other aspects of the NOS he noticed, but this elicitation does not lead to an answer. She then calls on two other students, which also met with silence. After creating a NOS learning opportunity through asking what NOS might be associated with the Mpemba effect example, Dr. Kim then decides to tell the students what other NOS aspects could be drawn from the story, such as the relationship of scientific knowledge and evidence, the advancement of experimental techniques leading to new discoveries, and the cumulative nature of scientific development. In this episode, we observe Dr. Kim's deliberate effort to use the presentation material to illuminate alternative aspects of NOS, even though it turned out to be irrelevant to the original learning goal (i.e., paradigm shift). From a formative assessment perspective, the lack of student responses functioned as diagnostic evidence that her initial question had not elicited understanding or engagement with the intended NOS ideas. Recognizing this, Dr. Kim adjusted her approach by providing direct instruction instead, shifting from elicitation to explanation in response to students' demonstrated difficulties.

In another episode, after Group 2 presented on plate tectonics and continental drift theory, Dr. Kim made another attempt to discuss other NOS aspects that could be said about the story. They presented what plate tectonics is, how the idea first appeared and became a paradigm, and what the benefits and limitations of plate tectonics are as an explanatory concept. At the end of their presentation, they revisited the definition of a paradigm shift from the previous lesson, and that plate tectonics meets the definition. When the presentation ended, Dr. Kim first recognized how the group connected the example and the paradigm shift, and then directed the class's attention toward NOS, specifically the tentative nature of science (Table 8).

She praises Bin's group for clearly linking continental drift to the paradigm shift and then shifts to another NOS aspect—the tentative nature of science (Line 1). After 7 s of silence (Line 2), she senses students' difficulty and recalls the earlier "electricity" lesson to prompt memory, reformulating her question into a fill-in-the-blank style ("Something leads to something") and offering a hint about science and technology (Line 3). When this still draws no response after another 5 s (Line 4), she simplifies further, asking whether limitations of plate tectonics can be overcome by advances in exploration technology (Line 6). Once students agree (Line 7), she generalizes the point: science depends on technology, and advances in one drive progress in the other. Apologizing for explaining rather than eliciting ideas, she acknowledges the missed opportunity for student reasoning (Line 8). From an ESRU perspective, this exchange shows incomplete formative cycles—her repeated prompts failed to generate student input, unlike earlier, more dialogic interactions that supported the NOS learning goal.

Collectively, her strategy to use the student presentations to shed light on NOS aspects other than paradigm shift was in the form of questioning, but none of these attempts led to her students answering her question. She gave hints and changed the question into the blank filling to make the question easier for students, but these efforts still did not lead to dialogic cycles of conversations. She ended up providing the additional NOS aspects on her own and touched on various epistemic and social aspects ranging from the tentativeness of science to other epistemic and social aspects of scientific knowledge. Again, this feedback by Dr. Kim was mostly impromptu and on-the-fly rather than planned, as she did not know the presentation contents in advance and had not taught the same lesson before.

5.5 | Synthesis: Building Connections Through Formative Assessment

Overall, Dr. Kim's enactment of the paradigm shift activity created a space for formative assessment of students' NOS learning through a range of instructional moves. To synthesize the analysis across the six group discussions, Table 9 provides an overview of which strategies were observed in each whole-class discussion. This summary highlights both the range and unevenness in how different strategies were enacted and responded to by students. As explained earlier, moves within "mapping historical examples onto NOS" and "focusing the discussion on NOS" often led to multi-turn discursive exchanges, suggesting that students became increasingly capable of identifying explicit connections between historical examples and NOS ideas, especially when scaffolded by teacher prompts or peer questions. In contrast, "exploring diverse NOS aspects" beyond the paradigm shift proved more challenging to produce multi-step ESRU cycles, although this was observed in five of the six whole-class discussions. Although Dr. Kim made deliberate attempts to elicit connections to additional NOS aspects, these questions were met with silence, and she ultimately provided the interpretations herself. This suggests that while students could engage with a single NOS idea in context, they struggled when asked to engage in more divergent thinking, possibly due to a lack of preparation or high-level fluency in NOS, especially without prior modeling or structured support.

Figure 1 is a graphical summary constructed from the analysis of Dr. Kim's formative assessment strategies that allowed for explicit-reflective NOS instruction (also see Table 4). It illustrates how her orchestration of NOS formative assessment facilitated connections between the central NOS learning goal with four related elements in the boxes in the corners, namely content knowledge, history of science, multiple historical examples, and other NOS aspects. These connections manifested in four themes—multicontextualizing NOS, mapping examples to NOS, focusing discussion on NOS, and exploring different aspects of the NOS.

TABLE 8 | Excerpt illustrating how Dr. Kim tried to relate the example to other aspects of NOS.

	Speaker	Utterance
1	Dr. Kim	Alright, well done to Bin's group who presented on continental drift. Next, let's have Q&A again. Seems like I can't give you much time. Just a minute. Let's have one question. You really explained well on the paradigm shift, and connected it well with continental drift theory, so I want to clap for that. *When we say the nature of science, we tend to talk about the tentative nature as a representative example. Anyone has a question related to that?*
2		(Silence, 7 s)
3	Dr. Kim	Is it difficult? Do you need another summary? When we studied with electricity earlier, don't you recall anything? Something leads to something. (Silence, 3 s) The fill-in-the-blank that you like. All blanks are for filling. Hint, the development of science and technology ...
4		(Silence, 5 s)
5		(Class laughs)
6	Dr. Kim	So, the limitations of plate tectonics, could be replaced [*sic: overcome*] through the advances in exploration techniques, right?
7	Class	Yes.
8	Dr. Kim	*So, relating to the tentative nature of science, science relies on technology, and one advance in science and technology can lead to another. We can say these things.* Guys, I'm really sorry for doing this summary myself. Well done.

TABLE 9 | Dr. Kim's formative assessment strategies observed in whole-class discussions.

Theme	Code	Group 1 (DNA)	Group 2 (plate tectonics)	Group 3 (heliocentric theory)	Group 4 (Mpemba effect)	Group 5 (Big bang theory)	Group 6 (vaccines)
Multicontextualizing NOS	Uses multiple historical episodes to teach paradigm shift	N/A[a]					
Mapping examples onto NOS	Clarifies learning goal	●					
	Explains a NOS concept	●					
	Helps relate the example to the paradigm shift	●			●	●	●
	Gives a compliment	●	●				
Focusing the discussion on NOS	Delivers relevant content knowledge			●		●	
	Clarifies learning goal			●			
	States expectations about content knowledge			●		●	
	Shuts down detailed content questions			●		●	
Exploring diverse NOS aspects	Explains a NOS concept		●				
	Helps relating the example to other NOS aspects	●	●	●	●		
	Connects to previous learning		●				

[a]The multicontextualization of NOS occurred during the lesson design process and during her concluding remark (see Section 5.1) so is not represented in this table.

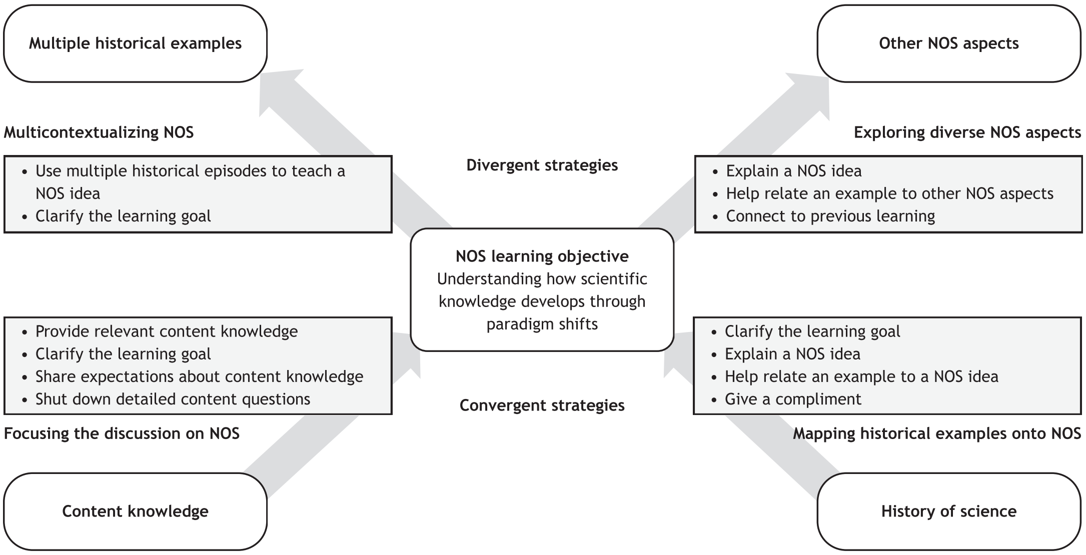

FIGURE 1 | Dr. Kim's formative assessment strategies that allowed for explicit-reflective NOS instruction.

We propose a categorization of the four identified strategies into *convergent* and *divergent* approaches to explicit-reflective NOS instruction. Convergent strategies are the teacher's conscious efforts to direct the focus to the target NOS idea instead of scientific and historical details while making connections with these components. Divergent strategies are efforts to promote learning a single NOS in multiple contexts or learning multiple NOS aspects from a single historical case. The ways she dealt with the background content knowledge and the historical details had a common aim: keeping the discussion focused on the key NOS learning goal, and preventing these contributing elements from becoming the focus of the lesson. However, since she believed that NOS should be understood in context rather than isolated statements, her feedback was concentrated on keeping the content knowledge and the historical context at an appropriate level. In this sense, these formative assessment practices represent Dr. Kim's effort to make the lesson converge on the learning goal and not be centered around the content knowledge or the historical details.

The other two themes, multicontextualizing and linking the example to other NOS aspects are divergent, in the sense that they have the potential to offer broader NOS understanding compared to when only one NOS is delivered using a single context. Considering the paradigm shift in multiple contexts and considering other NOS aspects of the same story requires students to not only have a NOS understanding but also creative thinking skills that connect their content knowledge, historical episodes and NOS aspects. This can pose challenges to both students and the teacher, particularly when the lesson heavily relies on student activities and participation. When the teacher attempted a divergent discussion about the NOS, the ESRU cycles were mostly incomplete (Table 9) due to the silences after teacher questions, possibly because the students did not have enough background knowledge that was necessary to build external links to NOS aspects or other historical examples other than what had been already covered in the lesson. Therefore, it is not surprising that despite the teacher's efforts, the students struggled in these areas (e.g., presenting examples that were not paradigm shifts and failing to respond to Dr. Kim's elicitation). This may indicate the need for the teacher to focus on a narrower range of content knowledge areas rather than opening up to all possible areas, or to align NOS learning more closely with the content knowledge curriculum.

Crucially, apart from the opportunities, her classroom episodes also revealed the persistent challenges that arise in both convergent and divergent processes, highlighting the difficulty of this work (Figure 2). Although studies have previously suggested a range of practical difficulties that can arise when teaching the NOS (Bartos and Lederman 2014; Cullinane and Erduran 2023; Hanuscin et al. 2011; Lederman 1999), the formative assessment conversations in Dr. Kim's classroom illustrate how those challenges can unfold for different instructional moves, how they are mutually related, and how the teacher reacted to those challenges. For example, episodes that showed the tension between "too much" and "too little" content knowledge suggest the difficulty of dealing with the content knowledge demand for effective learning of the NOS. All four practices Dr. Kim used to facilitate students' NOS learning were not without challenges, some of which stemmed from pedagogical considerations (e.g., the level of content knowledge required to understand the NOS), and some others inherent in NOS (e.g., the complexity of matching historical examples to NOS). While these are not an exhaustive list of challenges, they provide insights that can be incorporated into teacher education programs and instructional guidelines to help science teachers prepare to seize teachable moments for NOS and respond to potential challenges throughout the planning and delivery of instruction.

Given that the presentations and whole-class discussions spanned 2 weeks, it is worth noting that peer interactions became more sophisticated in the second week. This development appeared to stem from Dr. Kim's formative guidance: she increasingly prompted students to frame their questions around the NOS focus and modeled how to evaluate peers' reasoning against the learning goals. In this way, what initially began as general peer commentary evolved into more targeted formative

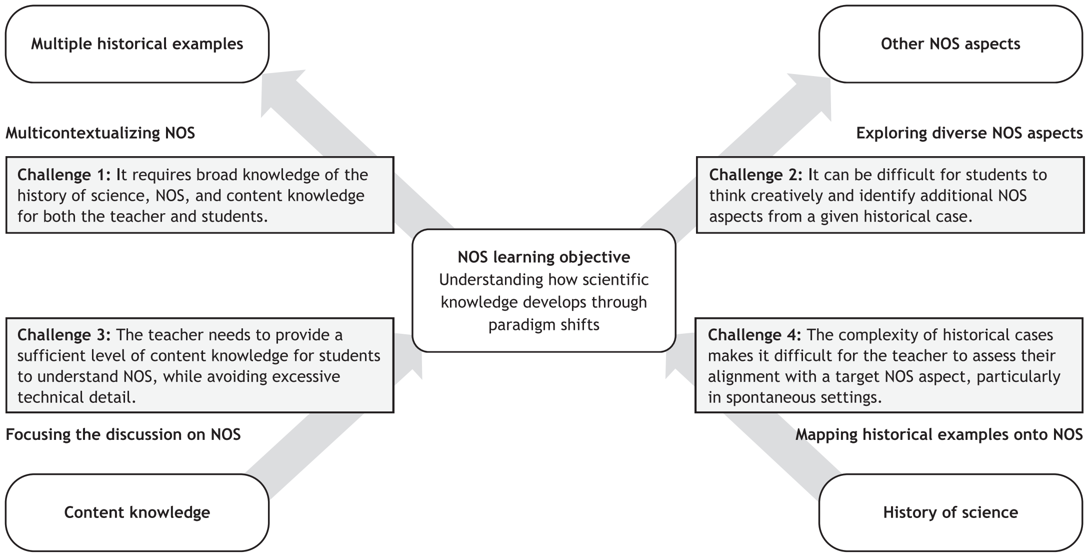

FIGURE 2 | Challenges of formative assessment of NOS observed in Dr. Kim's lesson.

assessment conversations. This shift illustrates how the teacher's responsive orchestration of peer assessment helped to deepen NOS-related dialogue, providing early evidence that the opportunities she created were beginning to translate into learning.

6 | Discussion

6.1 | Formative Assessment as a Lens to Analyze NOS Instructional Practice

In this study, we investigated the instructional strategies found in an experienced science teacher's lesson on the development of science. Specifically, we illustrated how Dr. Kim used assessment tasks, questioning, feedback, self-assessment and peer assessment to elicit, collect and act on evidence of students' NOS learning. Some of these practices were planned, but many were on-the-fly formative assessments. Dr. Kim contrasts with science teachers in previous reports in the literature, whose use of classroom assessment was limited when teaching the NOS (Hanuscin et al. 2011; Mesci et al. 2020; Ryder and Leach 2008). Bringing together theories of NOS instruction and those of formative assessment, we showed how analyzing NOS instruction through the lens of formative assessment can reveal the complex interplay of opportunities and difficulties inherent in NOS instruction through building connections between various components of NOS learning.

Previous research has advocated for the intentional teaching and assessment of NOS, with some studies pointing to teachers' limited use of classroom assessment when teaching NOS (Abd-El-Khalick et al. 1998; Hanuscin et al. 2011). Our analysis advances the understanding of this complex practice by using formative assessment as a micro-analytic lens to reveal how and to what extent an experienced teacher enacts the explicit-reflective NOS instruction in real time. By dissecting the discursive moves, our findings move beyond documenting the importance of FA to empirically demonstrating the specific strategies required to facilitate explicit and reflective NOS learning.

Overall, Dr. Kim's lessons showed how and to what extent NOS learning can be facilitated by the teacher's design and use of instructional tasks (formal) as well as questioning and feedback (informal), and how such formative assessments can be instrumental in "making connections" across content knowledge, the NOS and history of science, which has been stressed as a key element of NOS learning recently (Mulvey et al. 2021; Peters-Burton et al. 2019). Our analysis illustrates that formal and informal formative assessment can afford various opportunities to support explicit and reflective NOS learning: We observed positive indicators of student learning, such as an increase in multi-turn assessment conversations and student-initiated peer feedback becoming more frequent and critical. However, our study demonstrates that this process is not without considerable challenges and complexities. Formative assessment requires the teacher to have extensive knowledge of all many domains to enact prompt and responsive formative assessment, and build connections, throughout the instruction. Indeed, our analysis suggests that even for a highly experienced and knowledgeable teacher, enacting responsive formative assessment for NOS is a complex and challenging endeavor.

6.2 | Topic-Specificity of Formative Assessment and the Treatment of Content Knowledge

Research on formative assessment points to the need to consider the specificity of the subject and topic being taught (Black et al. 2003). As discussed earlier, NOS has distinct features in its curriculum content and necessitates the employment of different teaching approaches characterized as explicit and reflective instruction. The epistemic and reflective nature of the NOS makes it less amenable to standardized responses, which can complicate teachers' efforts to elicit and respond to student thinking (Brock and Park 2024). These considerations call for further examination of how the conditions for formative assessment to support students' learning come into play in the context of the NOS, which is open-ended and requires high-level thinking. In this study, Dr. Kim's lessons suggest that orchestrating a whole-class discussion on the NOS requires the teacher to have a broad range of scientific knowledge and its historical context (including outside their subject specialism), as well as pedagogical skills to design assessment tasks and make in-the-moment pedagogical reasoning to collect evidence of students' NOS learning and take appropriate action. Some of these knowledge and skills were specific to teaching and assessing NOS (as opposed to teaching and assessing science), affirming the earlier proposition that being a good NOS teacher requires more than being a good science teacher (Mesci et al. 2020).

6.3 | The Need for Empirically Grounded NOS Assessment Tasks and Learning Progressions

As we already pointed out earlier, despite their proven validity, using research instruments for classroom use has limitations in terms of meeting the needs of teachers and students. NOS researchers have emphasized using interactive assessment tasks to teach the NOS (Allchin et al. 2014; McComas 2020; Park, Wu, and Erduran 2020), where feedback and conversation occur more frequently than in teacher-centered direct instruction. Developing high-quality NOS assessment tasks and making them available to teachers will help to reduce the challenges identified in our study, such as the amount of background knowledge about scientific and historical details requisite for teachers to facilitate meaningful NOS formative assessment in the classroom. The assessment task that Dr. Kim designed and used was able to support learning in various ways, provided that the teacher had sufficient NOS knowledge and skills to make connections between content knowledge, NOS ideas, and the history of science. This can also mean that using formative assessment might be challenging for teachers with less experience with diverse content areas of science or NOS. NOS assessment frameworks and tasks informed by research on learning progressions (Alonzo 2011; Black et al. 2011; Pellegrino 2022) and a "developmental framework" that involves a progression of NOS benchmarks spanning from primary schools to universities and teacher education programs (Abd-El-Khalick 2014) might be of help.

6.4 | Understanding and Supporting Teachers' NOS Formative Assessment

Like other instructional practices, teachers need time to learn to use formative assessment effectively to support learning (Furtak 2012), and research suggests that formative assessment is hard to enact for science teachers, even with support (Furtak et al. 2008; Sezen-Barrie and Kelly 2017). Our analysis of Dr. Kim's practice and the associated challenges suggests that classroom assessment of NOS learning requires a set of skills and knowledge for teachers, some of which do not overlap with those skills and knowledge needed for formative assessment of science. This suggests that interventions and professional development programs focused on the NOS can include classroom assessment more explicitly, including task design and use, and informal assessment conversations, based on empirical evidence. Based on the results of this study, to enhance teachers' capacity to orchestrate assessment conversations responsive to NOS learning goals, professional development programs aimed at improving NOS assessment expertise could integrate several key elements: (a) opportunities to co-design and trial NOS assessment tasks in specific curricular contexts; (b) analysis of classroom videos to develop teachers' noticing and interpretation skills for NOS-related student thinking; and (c) structured reflection on how NOS ideas manifest in both student talk and artifacts (e.g., posters, reports, and presentations).

Given the limitations of single case study design, how NOS formative assessment practices can be influenced by teacher characteristics will need further investigation. Possible variations in practices and challenges of NOS formative assessment can be explored by including more teachers from different backgrounds in terms of subject specialism, NOS knowledge and pedagogical orientations. At the same time, as our analysis focused on whole-class interactions, how the teacher attends to students' individual experiences, aspirations and identities (Furtak and Lee 2023) in the planning and use of formative assessment of the NOS was not evident in our study. Future research with a focus on equity and justice in formative assessment will be necessary to provide further insights.

6.5 | Limitations

As an experienced science teacher with a strong personal interest in NOS and a deep understanding of its pedagogy, Dr. Kim was uniquely positioned to implement formative assessment strategies in NOS instruction. However, her ability to do so does not guarantee that other teachers with less experience or familiarity with the NOS would achieve similar outcomes. Similarly, although our results illustrate "what is possible" in the interplay between NOS and formative assessment, they are not an exhaustive account of how formative assessment can support explicit-reflective NOS instruction. In addition, it is important to recognize that this study was conducted in a naturalistic classroom setting without external input or intervention by the researchers. A more controlled intervention study that includes professional development on formative assessment for NOS may provide deeper insights into the opportunities and barriers associated with implementing these practices.

By focusing on a single, repeated pedagogical design, we were able to reveal a unique, fine-grained view into how a teacher and students negotiate meaning and epistemic goals within the constraints and opportunities of a single, coherent pedagogical approach; however, we acknowledge that this analytical approach does not capture the full repertoire of Dr. Kim's potential NOS assessment strategies. Furthermore, given that the focus on textual information does not capture students' meaning making through different modes of representations, such as pictures and gestures, the findings about students' understanding of NOS are limited. Recent research is beginning to address such aspects of NOS learning (Cheung et al. 2025) and future studies can also further understanding of how multimodality can be addressed in science teachers' learning of NOS teaching (Barak et al. 2023; Erduran and Kaya 2018). Finally, the study's context also warrants consideration. It was conducted under a new national curriculum that explicitly emphasized NOS as both a learning goal and an assessable component. This suggests that the effective enactment of NOS teaching is influenced not only by teacher expertise but also by curriculum design and policy context. Future research should explore how curriculum frameworks and professional development initiatives can collectively support teachers in successfully integrating NOS and formative assessment in their classrooms.

7 | Conclusion

Our analysis of Dr. Kim's NOS lessons illustrated the teacher's attempt to use a planned formative assessment activity and formative assessment conversations to make students' NOS learning visible in a whole class discussion setting. The episodes exemplified what teachers' use of questioning, noticing, feedback and peer assessment can look like in the context of the NOS. They demonstrate the value of formative assessment as a useful lens for improving research on NOS teaching and as an important avenue for future research. While this study did not use pre/post assessments to evaluate changes in students' NOS understanding, student interactions provide indirect evidence of increasing epistemic fluency, suggesting the ongoing influence of formative assessment in shaping learning opportunities over time. Given the limited empirical attention to formative assessment in NOS classroom practice, this study contributes to the literature by offering a detailed illustration of the complexities an experienced teacher faced when enacting formative assessment in NOS teaching, highlighting both the potential pedagogical moves and the inherent difficulties of this work. Our results also highlight the inherent complexities of assessing NOS. These complexities arise from the nature of NOS knowledge, which is intertwined with content knowledge, making its assessment challenging. Addressing these challenges requires further research into classroom practices to identify and refine pedagogical strategies for navigating the barriers associated with NOS formative assessment in science classrooms.

Acknowledgments

Wonyong Park would like to thank the Ilju Academy and Culture Foundation and the Department of Education at the University of Oxford, for scholarships to support his doctoral studies, from which this work was derived.

Data Availability Statement

Research data are not shared.

Endnotes

[1] The molecular structure of water does not appear in the national curriculum until 11th-grade Chemistry (an elective subject), so most students in the class had not learned about this in school unless they knew about it from other sources.

References

Abd-El-Khalick, F., and V. L. Akerson. 2004. "Learning as Conceptual Change: Factors Mediating the Development of Preservice Elementary Teachers' Views of Nature of Science." *Science Education* 88, no. 5: 785–810.

Abd-El-Khalick, F., R. L. Bell, and N. G. Lederman. 1998. "The Nature of Science and Instructional Practice: Making the Unnatural Natural." *Science Education* 82, no. 4: 417–436.

Abd-El-Khalick, F. 2014. "The Evolving Landscape Related to Assessment of Nature of Science." In *Handbook of Research on Science Education*, edited by N. G. Lederman and S. K. Abell, 621–650. Routledge.

Abd-El-khalick, F., and N. G. Lederman. 2000. "Improving Science Teachers' Conceptions of Nature of Science: A Critical Review of the Literature." *International Journal of Science Education* 22, no. 7: 665–701.

Abd-El-Khalick, F. 2012. "Nature of Science in Science Education: Toward a Coherent Framework for Synergistic Research and Development." In *Second International Handbook of Science Education*, edited by B. Fraser, K. Tobin, and C. J. McRobbie, 1041–1060. Springer.

Abd-El-Khalick, F. 2013. "Teaching With and About Nature of Science, and Science Teacher Knowledge Domains." *Science & Education* 22, no. 9: 2087–2107.

Abd-El-Khalick, F., and N. G. Lederman. 2023. "Research on Teaching, Learning, and Assessment of Nature of Science." In *Handbook of Research on Science Education*, edited by N. G. Lederman, D. L. Zeidler, and J. S. Lederman, vol. 3, 850–989. Routledge.

Abd-El-Khalick, F., R. Summers, J. L. Brunner, J. Belarmino, and J. Myers. 2024. "Development of VAScoR: A Rubric to Qualify and Score Responses to the Views of Nature of Science (VNOS) Questionnaire." *Journal of Research in Science Teaching* 61, no. 7: 1641–1688.

Akerson, V. L., F. Abd-El-Khalick, and N. G. Lederman. 2000. "Influence of a Reflective Explicit Activity-Based Approach on Elementary Teachers' Conceptions of Nature of Science." *Journal of Research in Science Teaching* 37, no. 4: 295–317.

Allchin, D. 2011. "Evaluating Knowledge of the Nature of (Whole) Science." *Science Education* 95, no. 3: 518–542.

Allchin, D. 2013. *Teaching the Nature of Science: Perspective & Resources*. SHiPs Education Press.

Allchin, D., H. M. Andersen, and K. Nielsen. 2014. "Complementary Approaches to Teaching Nature of Science: Integrating Student Inquiry, Historical Cases, and Contemporary Cases in Classroom Practice." *Science Education* 98, no. 3: 461–486.

Alonzo, A. C. 2011. "Learning Progressions That Support Formative Assessment Practices." *Measurement: Interdisciplinary Research and Perspectives* 9, no. 2–3: 124–129.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. 2014. *Standards for Educational and Psychological Testing*. American Educational Research Association.

Barak, M., T. Yachin, and S. Erduran. 2023. "Tracing Preservice Teachers' Understanding of Nature of Science Through Their Drawings and Writing." *Research in Science Education* 53: 507–523. https://doi.org/10.1007/s11165-022-10069-3.

Bartos, S. A., and N. G. Lederman. 2014. "Teachers' Knowledge Structures for Nature of Science and Scientific Inquiry: Conceptions and Classroom Practice." *Journal of Research in Science Teaching* 51, no. 9: 1150–1184.

Bell, B. 2007. "Classroom Assessment of Science Learning." In *Handbook of Research on Science Education*, edited by S. K. Abell and N. G. Lederman, 1st ed., 965–1005. Lawrence Erlbaum Associates.

Bell, B., and B. Cowie. 2001. "The Characteristics of Formative Assessment in Science Education." *Science Education* 85, no. 5: 536–553.

Bell, R. L. 2008. *Teaching the Nature of Science Through Process Skills: Activities for Grades 3–8*. Pearson.

Birks, M., Y. Chapman, and K. Francis. 2008. "Memoing in Qualitative Research: Probing Data and Processes." *Journal of Research in Nursing* 13, no. 1: 68–75.

Black, P., C. Harrison, C. Lee, B. Marshall, and D. Wiliam. 2003. *Assessment for Learning: Putting It Into Practice*. Open University Press.

Black, P., and D. Wiliam. 2009. "Developing the Theory of Formative Assessment." *Educational Assessment, Evaluation and Accountability* 21, no. 1: 5–31.

Black, P., M. Wilson, and S. Y. Yao. 2011. "Road Maps for Learning: A Guide to the Navigation of Learning Progressions." *Measurement: Interdisciplinary Research and Perspectives* 9, no. 2–3: 71–123.

Black, P. J. 1998. *Testing, Friend or Foe?: The Theory and Practice of Assessment and Testing*. Falmer.

Brinkmann, S., and S. Kvale. 2015. *Interviews: Learning the Craft of Qualitative Research Interviewing*. 3rd ed. Sage.

Brock, R., and W. Park. 2024. "Distinguishing Nature of Science Beliefs, Knowledge and Understandings: Towards Clarity and Coherence in Educational Goals Related to the Nature of Science." *Science & Education* 33, no. 3: 495–516.

Brock, R., and K. S. Taber. 2019. "I'm Sad That It's Gone: A Case Study of Teachers' Views on Teaching the Nature of Science at Key Stage 4." *School Science Review* 100, no. 373: 29–34.

Cheung, K. K. C., and S. Erduran. 2025. "Designing an Analytical Framework to Investigate Students' Multimodal Representations of Scientific Practices and Methods." *Research in Science Education*. https://doi.org/10.1007/s11165-025-10248-y.

Cheung, K. K. C., A. Oancea, and S. Erduran. 2025. "Students' Meaning-Making of Nature of Science: Interaction Between Visual, Verbal, and Written Modes of Representation." *Science Education* 109: 480–505.

Cizek, G. J., H. L. Andrade, and R. E. Bennett. 2019. "Formative Assessment: History, Definition, and Progress." In *Handbook of Formative Assessment in the Disciplines*, edited by H. L. Andrade, R. E. Bennett, and G. J. Cizek, 3–19. Routledge.

Cullinane, A., and S. Erduran. 2023. "Nature of Science in Preservice Science Teacher Education—Case Studies of Irish Pre-Service Science Teachers." *Journal of Science Teacher Education* 34, no. 2: 201–223.

Deng, F., D. T. Chen, C. C. Tsai, and C. S. Chai. 2011. "Students' Views of the Nature of Science: A Critical Review of Research." *Science Education* 95, no. 6: 961–999.

Deng, F., Z. Shi, P. Tian, and D. Sun. 2025. "Explicit-Reflective Teaching and Students' Understanding of Nature of Science: A Meta-Analysis." *Research in Science Education*. https://doi.org/10.1007/s11165-025-10287-5.

Deverel-Rico, C., and E. M. Furtak. 2025. "How Can Features of Classroom Activity Systems Center Students' Ideas and Experiences in Formative Assessment? A Study of Sensemaking and Answer-Making Frames." *Educational Assessment* 30, no. 2: 91–114.

Duschl, R. A., and D. H. Gitomer. 1997. "Strategies and Challenges to Changing the Focus of Assessment and Instruction in Science Classrooms." *Educational Assessment* 4, no. 1: 37–73.

Empson, S. B., and V. R. Jacobs. 2008. "Learning to Listen to Children's Mathematics." In *International Handbook of Mathematics Teacher Education: Tools and Processes in Mathematics Teacher Education*, edited by D. Tirosh and T. Wood, vol. II, 257–281. Sense Publishers.

Erduran, S., and Z. R. Dagher. 2014. *Reconceptualizing the Nature of Science for Science Education: Scientific Knowledge, Practices and Other Family Categories*. Springer.

Erduran, S., and E. Kaya. 2018. "Drawing Nature of Science in Pre-Service Science Teacher Education: Epistemic Insight Through Visual Representations." *Research in Science Education* 48: 1133–1149. https://doi.org/10.1007/s11165-018-9773-0.

Ferreira, S., and A. M. Morais. 2013. "The Nature of Science in Science Curricula: Methods and Concepts of Analysis." *International Journal of Science Education* 35, no. 16: 2670–2691.

Furtak, E. M. 2012. "Linking a Learning Progression for Natural Selection to Teachers' Enactment of Formative Assessment." *Journal of Research in Science Teaching* 49, no. 9: 1181–1210.

Furtak, E. M. 2023. *Formative Assessment for 3D Science Learning: Supporting Ambitious and Equitable Instruction*. Teachers College Press.

Furtak, E. M., S. C. Heredia, and D. Morrison. 2019. "Formative Assessment in Science Education: Mapping a Shifting Terrain." In *Handbook of Formative Assessment in the Disciplines*, edited by H. L. Andrade, R. E. Bennett, and G. J. Cizek, 97–125. Routledge.

Furtak, E. M., K. Kiemer, R. K. Circi, et al. 2016. "Teachers' Formative Assessment Abilities and Their Relationship to Student Learning: Findings From a Four-Year Intervention Study." *Instructional Science* 44: 267–291.

Furtak, E. M., and O. Lee. 2023. "Equity and Justice in Classroom Assessment of STEM Learning." In *Classroom-Based STEM Assessment: Contemporary Issues and Perspectives*, edited by C. J. Harris, E. Wiebe, S. Grover, and J. W. Pellegrino, 69–85. Community for Advancing Discovery Research in Education (CADRE). Education Development Center, Inc.

Furtak, E. M., M. A. Ruiz-Primo, J. T. Shemwell, et al. 2008. "On the Fidelity of Implementing Embedded Formative Assessments and Its Relation to Student Learning." *Applied Measurement in Education* 21, no. 4: 360–389.

Furtak, E. M., J. Thompson, and E. A. van Es. 2016. "Formative Assessment and Noticing: Toward a Synthesized Framework for Attending and Responding During Instruction." In *Annual Meeting of the American Educational Research Association, Washington, DC*, vol. 9. American Educational Research Association.

Glaser, B., and A. Strauss. 2017. *Discovery of Grounded Theory: Strategies for Qualitative Research*. Routledge.

Hanuscin, D., M. Khajeloo, and B. C. Herman. 2020. "Considering the Classroom Assessment of Nature of Science." In *Nature of Science in Science Instruction: Rationales and Strategies*, edited by W. F. McComas, 409–423. Springer.

Hanuscin, D. L., M. H. Lee, and V. L. Akerson. 2011. "Elementary Teachers' Pedagogical Content Knowledge for Teaching the Nature of Science." *Science Education* 95, no. 1: 145–167.

Heritage, M. 2013. "Gathering Evidence of Student Understanding." In *SAGE Handbook of Research on Classroom Assessment*, edited by J. H. McMillan, 179–196. Sage.

Kampourakis, K. 2016. "The "General Aspects" Conceptualization as a Pragmatic and Effective Means to Introducing Students to Nature of Science." *Journal of Research in Science Teaching* 53, no. 5: 667–682.

Kang, H., and E. M. Furtak. 2021. "Learning Theory, Classroom Assessment, and Equity." *Educational Measurement: Issues and Practice* 40, no. 3: 73–82.

Khishfe, R., and F. Abd-El-Khalick. 2002. "Influence of Explicit and Reflective Versus Implicit Inquiry-Oriented Instruction on Sixth Graders' Views of Nature of Science." *Journal of Research in Science Teaching* 39, no. 7: 551–578.

Khishfe, R. 2023. "Connected Learning: An Approach for Teaching Nature of Science Aspects and Argumentation Components." *Journal of Research in Science Teaching* 60, no. 10: 2180–2215.

Leach, J., R. Millar, J. Ryder, and M.-G. Séré. 2000. "Epistemological Understanding in Science Learning: The Consistency of Representations Across Contexts." *Learning and Instruction* 10: 497–527.

Leahy, S., C. Lyon, M. Thompson, and D. Wiliam. 2005. "Classroom Assessment: Minute by Minute, Day by Day." *Educational Leadership* 63, no. 3: 18–24.

Lederman, N. G., and F. Abd-El-Khalick. 1998. "Avoiding de-Natured Science: Activities That Promote Understandings of the Nature of Science." In *The Nature of Science in Science Education: Rationales and Strategies*, edited by W. F. McComas, 83–126. Springer.

Lederman, N. G. 1999. "Teachers' Understanding of the Nature of Science and Classroom Practice: Factors That Facilitate or Impede the Relationship." *Journal of Research in Science Teaching* 36, no. 8: 916–929.

Lederman, N. G., F. Abd-El-Khalick, R. L. Bell, and R. S. Schwartz. 2002. "Views of Nature of Science Questionnaire: Toward Valid and Meaningful Assessment of Learners' Conceptions of Nature of Science." *Journal of Research in Science Teaching* 39, no. 6: 497–521.

Lederman, N. G., R. S. Schwartz, F. Abd-El-Khalick, and R. L. Bell. 2001. "Pre-Service Teachers' Understanding and Teaching of Nature of Science: An Intervention Study." *Canadian Journal of Science, Mathematics and Technology Education* 1, no. 2: 135–160.

Lederman, N. G. 2007. "Nature of Science: Past, Present, and Future." In *Handbook of Research on Science Education*, edited by S. K. Abell and N. G. Lederman, 831–880. Erlbaum.

Martone, A., and S. G. Sireci. 2009. "Evaluating Alignment Between Curriculum, Assessment, and Instruction." *Review of Educational Research* 79, no. 4: 1332–1361.

McComas, W. F. 2020. "Nature of Science and Classroom Practice: A Review of the Literature With Implication for Effective NOS Instruction." In *Nature of Science in Science Instruction: Rationales and Strategies*, edited by W. F. McComas, 67–111. Springer.

McDonald, C. V. 2010. "The Influence of Explicit Nature of Science and Argumentation Instruction on Preservice Primary Teachers' Views of Nature of Science." *Journal of Research in Science Teaching* 47, no. 9: 1137–1164.

McMillan, J. H. 2010. "The Practical Implications of Educational Aims and Contexts for Formative Assessment." In *Handbook of Formative Assessment*, edited by H. L. Andrade and G. J. Cizek, 41–58. Routledge.

Mehan, H. 1982. "The Structure of Classroom Events and Their Consequences for Student Performance." In *Children in and out of School*, edited by P. Gilmore and A. A. Glatthorn. Center for Applied Linguistics.

Mercer, N., L. Dawes, and J. K. Staarman. 2009. "Dialogic Teaching in the Primary Science Classroom." *Language and Education* 23, no. 4: 353–369.

Mesci, G., R. S. Schwartz, and B. A. S. Pleasants. 2020. "Enabling Factors of Preservice Science Teachers' Pedagogical Content Knowledge for Nature of Science and Nature of Scientific Inquiry." *Science & Education* 29, no. 2: 263–297.

Ministry of Education. 2015. "National Curriculum: Science."

Morse, J. M. 1997. ""Perfectly Healthy, but Dead": The Myth of Inter-Rater Reliability." *Qualitative Health Research* 7, no. 4: 445–447.

Mpemba, E. B., and D. G. Osborne. 1969. "Cool?" *Physics Education* 4, no. 3: 172–175.

Mulvey, B. K., J. C. Parrish, J. W. Reid, J. Papa, and E. E. Peters-Burton. 2021. "Making Connections: Using Individual Epistemic Network Analysis to Extend the Value of Nature of Science Assessment." *Science & Education* 30, no. 3: 527–555.

National Research Council. 2001. *Classroom Assessment and the National Science Education Standards*. National Academies Press.

Nieminen, P., C. F. Correia, M. Hähkiöniemi, N. Serret, J. Viiri, and C. Harrison. 2016. "Formative Assessment in Inquiry-Based Science Education Using Interactions On-the-Fly." In *Paper Presented in NARST Annual International Conference, Baltimore, MD*. Springer.

Oliveira, G. 2023. "Developing a Codebook for Qualitative Data Analysis: Insights From a Study on Learning Transfer Between University and the Workplace." *International Journal of Research & Method in Education* 46, no. 3: 300–312.

Olson, J. K. 2018. "The Inclusion of the Nature of Science in Nine Recent International Science Education Standards Documents." *Science & Education* 27, no. 7–8: 637–660.

Panadero, E. 2016. "Is It Safe? Social, Interpersonal, and Human Effects of Peer Assessment: A Review and Future Directions." In *Handbook of Social and Human Conditions in Assessment*, edited by G. T. L. Brown and L. R. Harris, 247–266. Routledge.

Panadero, E., A. Jonsson, and J. Botella. 2017. "Effects of Self-Assessment on Self-Regulated Learning and Self-Efficacy: Four Meta-Analyses." *Educational Research Review* 22: 74–98.

Park, W., J.-Y. Wu, and S. Erduran. 2020. "The Nature of STEM Disciplines in the Science Education Standards Documents From the USA, Korea and Taiwan: Focusing on Disciplinary Aims, Values and Practices." *Science & Education* 29, no. 4: 899–927.

Park, W., S. Yang, and J. Song. 2020. "Eliciting Students' Understanding of Nature of Science With Text-Based Tasks: Insights From New Korean High School Textbooks." *International Journal of Science Education* 42, no. 3: 426–450.

Pellegrino, J. W. 2022. "A Learning Sciences Perspective on the Design and Use of Assessment in Education." In *The Cambridge Handbook of the Learning Sciences*, 238–258. Cambridge University Press.

Peters-Burton, E. E., and S. R. Burton. 2020. "The Use of Metacognitive Prompts to Foster Nature of Science Learning." In *Nature of Science in Science Instruction: Rationales and Strategies*, edited by W. F. McComas, 179–198. Springer.

Peters-Burton, E. E., J. C. Parrish, and B. K. Mulvey. 2019. "Extending the Utility of the Views of Nature of Science Assessment Through Epistemic Network Analysis." *Science & Education* 28, no. 9–10: 1027–1053.

Rached, E., and M. Grangeat. 2021. "French Teachers' Informal Formative Assessment in the Context of Inquiry-Based Learning." *International Journal of Science Education* 43, no. 3: 385–406.

Reinisch, B., and K. Fricke. 2022. "Broadening a Nature of Science Conceptualization: Using School Biology Textbooks to Differentiate the Family Resemblance Approach." *Science Education* 106, no. 6: 1375–1407.

Roberts, D. A., A. M. MacKinnon, and M. E. McCarthy. 1992. "Reasons for Giving Reasons: An Expert-Expert Clinical Analysis of Science Teaching for Non-Academic Students." *Qualitative Studies in Education* 5, no. 3: 227–243.

Ruiz-Primo, M. A., and E. M. Furtak. 2007. "Exploring Teachers' Informal Formative Assessment Practices and Students' Understanding in the Context of Scientific Inquiry." *Journal of Research in Science Teaching* 44, no. 1: 57–84.

Ruiz-Primo, M. A. 2011. "Informal Formative Assessment: The Role of Instructional Dialogues in Assessing Students' Learning." *Studies in Educational Evaluation* 37, no. 1: 15–24.

Ryder, J., and J. Leach. 2008. "Teaching About the Epistemology of Science in Upper Secondary Schools: An Analysis of Teachers' Classroom Talk." *Science & Education* 17, no. 2–3: 289–315. https://doi.org/10.1007/s11191-006-9007-0.

Sadler, D. R. 1989. "Formative Assessment and the Design of Instructional Systems." *Instructional Science* 18, no. 2: 119–144.

Sadler, T. D. 2014. "Assessment of Socio-Scientific Reasoning: Linking Progressive Aims of Science Education to the Realities of Modern

Education." In *Assessing Schools for Generation R (Responsibility)*, edited by M. P. Mueller, D. J. Tippins, and A. J. Stewart, 101–113. Springer.

Saldaña, J. 2013. *The Coding Manual for Qualitative Researchers*. Sage.

Schwartz, R. S., and N. G. Lederman. 2002. ""It's the Nature of the Beast": The Influence of Knowledge and Intentions on Learning and Teaching Nature of Science." *Journal of Research in Science Teaching* 39, no. 3: 205–236.

Sezen-Barrie, A., and G. J. Kelly. 2017. "From the Teacher's Eyes: Facilitating Teachers Noticings on Informal Formative Assessments (IFAs) and Exploring the Challenges to Effective Implementation." *International Journal of Science Education* 39, no. 2: 181–212.

Shavelson, R. J., D. B. Young, C. C. Ayala, et al. 2008. "On the Impact of Curriculum-Embedded Formative Assessment on Learning: A Collaboration Between Curriculum and Assessment Developers." *Applied Measurement in Education* 21, no. 4: 295–314.

Silverman, D. 2014. *Interpreting Qualitative Data*. 5th ed. Sage.

Sinclair, J., and M. Coulthard. 1975. *Towards an Analysis of Discourse*. Oxford University Press.

Tang, K.-S. 2021. *Discourse Strategies for Science Teaching and Learning: Research and Practice*. Routledge.

Tseng, S. C., and C. C. Tsai. 2007. "On-Line Peer Assessment and the Role of the Peer Feedback: A Study of High School Computer Course." *Computers & Education* 49, no. 4: 1161–1174.

VERBI Software. 2019. *MAXQDA 2020 [Computer Software]*. VERBI Software. http://maxqda.com/.

Wahbeh, N., and F. Abd-El-Khalick. 2014. "Revisiting the Translation of Nature of Science Understandings Into Instructional Practice: Teachers' Nature of Science Pedagogical Content Knowledge." *International Journal of Science Education* 36, no. 3: 425–466.

Wiliam, D. 2009. *Assessment for Learning: Why, What and How?* Institute of Education.

Wiliam, D. 2010. "An Integrative Summary of the Research Literature and Implications for a New Theory of Formative Assessment." In *Handbook of Formative Assessment*, edited by H. L. Andrade and G. J. Cizek, 18–40. Routledge.

Wong, S. L., J. Kwan, D. Hodson, and B. H. W. Yung. 2009. "Turning Crisis Into Opportunity: Nature of Science and Scientific Inquiry as Illustrated in the Scientific Research on Severe Acute Respiratory Syndrome." *Science & Education* 18, no. 1: 95–118.

Yeh, S. S. 2010. "Understanding and Addressing the Achievement Gap Through Individualized Instruction and Formative Assessment." *Assessment in Education: Principles, Policy & Practice* 17, no. 2: 169–182.

Yin, R. K. 2017. *Case Study Research and Applications*. 6th ed. Sage.