frontiers | Frontiers in Education

TYPE Hypothesis and Theory
PUBLISHED 03 September 2024
DOI 10.3389/feduc.2024.1452111

OPEN ACCESS

EDITED BY
Alfonso Garcia De La Vega,
Autonomous University of Madrid, Spain

REVIEWED BY
Colin Evers,
University of New South Wales, Australia
Isabel María Montaraz Moral,
Autonomous University of Madrid, Spain

*CORRESPONDENCE
Annelies Pieterman-Bos
✉ j.e.bos-11@umcutrecht.nl

RECEIVED 20 June 2024
ACCEPTED 19 August 2024
PUBLISHED 03 September 2024

CITATION
Pieterman-Bos A, Reincke CM, van de Schoot R and van Mil MHW (2024) Understanding epistemological notions underlying scientific language use: a multifaceted analysis framework. *Front. Educ.* 9:1452111. doi: 10.3389/feduc.2024.1452111

Understanding epistemological notions underlying scientific language use: a multifaceted analysis framework

Annelies Pieterman-Bos[1]*, Cathelijne M. Reincke[1], Rens van de Schoot[2] and Marc H. W. van Mil[1]

[1]Center of Education and Training, University Medical Center Utrecht, Utrecht, Netherlands, [2]Department of Methodology and Statistics, Faculty of Social and Behavioral Sciences, Utrecht University, Utrecht, Netherlands

Helping university students develop productive views of what characterizes good scientific research and scientific knowledge is an important objective of science education. However, many studies show that students' views of the nature of science (NOS) do not become more informed or even become less informed after learning activities that engage students in scientific practice. This could mean that the way we teach students to conduct scientific research might unintentionally strengthen or bring about uninformed views of NOS. In this article, we argue that scientific discourse (how language is used in science) might play a role in this relation between learning scientific inquiry and uninformed views of NOS. We argue that there could be implicit notions relating to NOS underlying scientific discourse that uninformed readers might translate to uninformed views of NOS. We call these implicit notions underlying scientific language use "epistemological notions." In this article, we further define this construct of epistemological notions, contrast it with explicit views of NOS and other related constructs, explain how we think epistemological notions might affect explicit views of NOS, and present a framework we developed to characterize them in scientific language use. It is a descriptive and interpretative analysis framework which combines, optimizes, and extends several text analysis methods, discourse analysis, and reflexive thematic analysis. We provide a guide to use the framework and point out quality criteria. We finish by advocating the framework for educational researchers interested in developing instructional interventions during which learning about science is combined with explicit reflection on NOS. Identifying epistemological notions in scientific language use could provide starting points for these activities by making explicit the translation from NOS understanding to scientific practice and vice versa.

KEYWORDS

nature of science, epistemological notions, epistemology, research method, scientific discourse, philosophy of science

1 Introduction

Helping university students develop productive views of what characterizes good scientific research and scientific knowledge is an important objective of science education. Many studies have been conducted to understand how students at different school levels view the nature, origin and limits of knowledge and the process of knowing (Erduran et al., 2019; Cobo

et al., 2022; Abd-El-Khalick and Lederman, 2023). Theories about knowledge are known as epistemologies. Science education and research attending to epistemology often focuses on students' views about a specific subset of characteristics of the nature of scientific knowledge and knowing. This subset of views is defined by science educators, and they are often called views of the nature of science (NOS) (e.g., Lederman, 1992). Most studies of students' NOS views focus on what students say their views of NOS are when this is explicitly discussed in class or for the purpose of educational research (Deng et al., 2011). Deng et al. (2011) call these views "professed views of the nature of science." Science educators aim to change students' uninformed views of NOS to informed views. To that end, there are three common types of NOS instructional interventions: explicit-reflective interventions where NOS is explicitly addressed in class, implicit interventions where NOS is implicitly taught through student participation in science, and a combination of participation in science and explicit reflection on NOS (Abd-El-Khalick and Lederman, 2023). Most studies using an implicit approach show that students' NOS views do not become more informed or even become less informed after mere experience with scientific practice, either through inquiry or exposition (Cobo et al., 2022; Metin Peten, 2022; Abd-El-Khalick and Lederman, 2023). This could mean that the way we teach students to conduct scientific research could unintentionally bring about or strengthen uninformed views of NOS. In this article, we argue that scientific discourse (how language is used in science) might play a role in this relation between learning scientific inquiry and uninformed views of NOS. For example, writing objectively about scientific observations (e.g., omitting first-person sentences), could give students the wrong impression that science starts with neutral observations, independent of the perspectives of the researcher doing these observations. In other words, there might be implicit notions relating to NOS (e.g., scientific observations are objective) underlying scientific discourse (e.g., omitting first-person sentences) that uninformed readers (e.g., students) might translate to uninformed views of NOS (e.g., science starts with neutral observation). While more informed readers (e.g., professional scientists) do not register these implicit notions relating to NOS or do not translate these to uninformed views. We call these implicit notions underlying scientific language use "epistemological notions." In this article, we further define this construct of epistemological notions, contrast it with explicit views of NOS and other related constructs, explain how we think epistemological notions might affect explicit views of NOS, and present a framework we developed to characterize them in scientific language use. We built on several previously published text analysis methods with similar but smaller purposes. We combined, optimized, and extended this work to construct a more comprehensive analysis framework: the epistemological notions analysis framework.

In addition, we emphasize the need for explicit reflection on NOS when teaching (about) scientific language use, since there is strong evidence that instructional interventions focused on learning about science or learning to do science need to be combined with explicit reflection on NOS for them to result in more informed views of NOS (Cobo et al., 2022; Metin Peten, 2022; Abd-El-Khalick and Lederman, 2023). We argue that identifying epistemological notions in scientific language use could provide starting points for learning activities in which we explicitly reflect with students on these epistemological notions and how they relate to views of NOS and, more broadly, epistemology.

The main aim of this article is to define our newly introduced concept of epistemological notions and to present an analysis framework to characterize them in scientific language use. First, we outline the theoretical background and development of the framework. Then, we present the framework as a six-phase process to guide epistemological notions analysis and point out quality criteria. Lastly, we discuss the value and applicability of the framework.

2 Theoretical background

Before we introduce the analysis framework, we provide some theoretical background for using the framework. We start by summarizing existing text analysis approaches. Next, we introduce and define the concept of epistemological notions. Then, we argue how they are related to explicit views of NOS and epistemology and describe how views of NOS are defined in literature. Lastly, we describe three theoretical and philosophical underpinnings of the analysis framework.

2.1 Review of existing text analysis approaches

The epistemological notions analysis framework is based on various existing text analysis approaches. Roughly, we can discern three ways to analyze academic texts that could provide insights into underlying epistemological notions: analyzing the structure, quality, and language of an argument. These approaches describe different aspects of how writers create arguments for scientific claims. We have combined them into the first building block of the epistemological notions analysis framework (descriptive text analysis). The other building blocks are discourse analysis (Gee, 2014) and reflexive thematic analysis (Braun and Clarke, 2006, 2021), discussed in Section 2.4. First, we summarize what aspects of the structure, quality, or language of an argument have been studied by others.

2.1.1 Structure of arguments

Concerning the structure of arguments, some influential research lines started with the work of Toulmin (1958) on argument structure and the work of Latour and Woolgar (1986) on the construction of scientific facts. Since Toulmin's structure is not directly applicable to more complex arguments, Science education researchers have extended the structure to academic arguments (Kelly et al., 1998; Kelly and Chen, 1999; Kelly and Takao, 2002; Takao and Kelly, 2003; Kelly et al., 2007; Jiménez-Aleixandre and Federico-Agraso, 2009). They studied both the structure and quality of arguments. Concerning the structure, they assessed the coordination of evidence across epistemic levels by reconstructing the argument structure based on its argument components: data inscription, evidence, justification, and claim (see Table 1 for an explanation of these argument components). These reconstructions of argument structures are subsequently used to further assess the quality of arguments.

2.1.2 Quality of arguments

To assess the quality of the arguments, previous studies evaluated the pertinence (relevance) of the evidence (Jiménez-Aleixandre and Federico-Agraso, 2009), the sufficiency of the evidence (Sandoval and Millwood, 2005; Kelly et al., 2007; Jiménez-Aleixandre and

TABLE 1 Argument aspects assessed in existing text analysis approaches.

Argument aspect	Code	Description
Argument components (Jiménez-Aleixandre and Federico-Agraso, 2009)	Data inscription	The tables and figures of the article. For ease of reading, we refer to data inscriptions as "table," "figure," or "tables and figures."
	Evidence	Sentences about the experimental results supporting the claims.
	Justification	Establishes the connection between a data inscription or evidence sentence and a claim (the equivalent of Toulmin's warrant). Justifications can be direct, being a simple justification in support of a claim, or subsequent, being both the endpoint for one argument (i.e., a claim) and a connection between evidence and claim in a second argument (i.e., a justification).
	Claim	The thesis the authors are seeking to demonstrate.
Coordination of the evidence across epistemic levels (Jiménez-Aleixandre and Federico-Agraso, 2009)	Sufficient/insufficient	Epistemic level relates to the degree of abstractness of knowledge claims, from grounded, low inference claims to progressively more general, theoretical claims. Argument structure succeeds from data inscription to evidence (sentence), to first level claim, to second level claim.
Pertinence of the evidence (Jiménez-Aleixandre and Federico-Agraso, 2009)	Pertinent/not pertinent	Is the evidence presented relevant for the claim?
Level of rhetorical reference to data inscriptions (Sandoval and Millwood, 2005)	Inclusion	A figure or table is included in the explanation without reference to the inscription in the text.
	Pointer	A non-descriptive reference to a figure or table (e.g., "See graph 1").
	Description	A summary or other description of the figure or table with no suggestion of its relation to a claim.
	Assertion	A sentence about a figure or table in which the figure or table is asserted to show or prove a claim, without an explanation as to how it does so.
	Interpretation	A sentence that explicitly relates specific features of an inscription to a claim.
Type of warrant used for a claim (Sandoval and Millwood, 2007)	Authority	Instances where a student explicitly states a source of authority or lack thereof (e.g., teacher, class, book).
	Causal	Warrants that refer to reasons based on a theoretical concept, or explanation of a theoretical concept.
	Empirical	Reasons citing some kind of empirical evidence or lack thereof (e.g., literature reference, data).
	Factual	Repeating of the original claim by using the exact same words, paraphrasing, or rephrasing.
Statement types (Latour and Woolgar, 1986)	Type 5	Statements corresponding to a taken-for-granted fact.
	Type 4	Statements explicitly addressing a "fact" as such.
	Type 3	Statements about other statements, using modalities and references.
	Type 2	Statements that contain modalities which draw attention to the circumstances of the evidence affecting the level of generality of the statement.
	Type 1	Statements that are conjectures or speculations.
Modality (Hyland, 1998; Plappert, 2019)	Hedge	Words used to decrease the certainty of a statement.
	Booster	Words used to increase the certainty of a statement.
	Implicature	Words used to speculate at the level of speaker meaning (what a writer implicates) while at the sentential level (what a writer writes) the statement has a higher level of certainty. Example: "Gene X is associated with disease Y." At the level of speaker meaning, it suggests that Y *might* be caused by A.

10.3389/feduc.2024.1452111

Federico-Agraso, 2009), the level of rhetorical reference to data inscriptions (Sandoval and Millwood, 2005), and the type of justifications used for claims (Sandoval and Millwood, 2007). Kelly et al. (2007) assessed sufficiency of the evidence and coordination across epistemic levels. This assessment covers different aspects, each on a scale from 0 (non-existent) to 4 (excellent). These aspects include the following: solvable and clearly stated thesis or question; multiplicity and plausibility of lines of reasoning; potential sufficiency of data; inclusion, identification, and description of appropriate data representations and their use to support an explanation; relevancy and clear identification of the relevancy of the data used; validity of inferences; convergence of lines of reasoning to a conclusion; and overall support of the thesis. Jiménez-Aleixandre and Federico-Agraso (2009) assessed pertinence, sufficiency of evidence, and coordination across epistemic levels. They did not distinguish different aspects but assessed all three as a single measure and as either adequate or not.

The arguments assessed by Sandoval and Millwood (2005) were of a slightly different type. They assessed a student assignment where high school students were asked to explain specific evolution-related questions through exploration of provided data. They assessed some aspects of argument quality based on criteria that depend on the strong delineation of the assignment, which are not applicable for the current study. However, Sandoval and Millwood do specify an aspect of quality of arguments that is relevant for the current study, the "level of rhetorical reference to data inscriptions." Through their analysis of the use of data inscriptions (i.e., figures and tables) in student explanations, they identified five levels of rhetorical use: inclusion, pointer, description, assertion, and interpretation (see Table 1).

In another study, Sandoval and Millwood (2007) characterized what type of warrants students provided for their scientific claims. They identified four types of warrants in interviews with students about their conceptions of the best way to convince someone of something in science: authority, causal, empirical, and factual (see Table 1). Sandoval and Millwood describe that they have compared student responses with their actual use of warrants in the essays, although they do not go into detail on how they did so. Nevertheless, these categories might also be recognizable in scientific language use.

2.1.3 Language of arguments

The last aspect of argumentation that could provide insight in the epistemological notions underlying scientific writing is the language used. The language used to communicate scientific claims can signal their perceived epistemic status (i.e., the degree of certainty awarded to knowledge claims). Three studies that are of importance for the current study are those of Latour and Woolgar (1986), Hyland (1998), and Plappert (2019). These studies are centered around the language of varying strengths of knowledge claims and based on philosophical and sociological studies of Bruno Latour, Thomas Kuhn, George Lakoff, and Paul Grice.

Kuhn (1970, p.127, 1982) already exemplifies that knowledge is not just conveyed but constituted by the use of language with his introduction of the incommensurability thesis. Latour and Woolgar (1986, p. 79) in addition, argue that the certainty of knowledge claims can be recognized in their linguistic structures, although there is no simple relationship between these structures and the level of certainty they express. Latour studied the extent to which some statements appeared more fact-like than others. He recognized five statement types that express different levels of certainty of a "fact" (see Table 1). However, the form of a statement does not always directly correspond with its recognized level of certainty. For example, the inclusion of a reference might, on the one hand, detract from the certainty of a statement because it denotes the involvement of human agency in its production. On the other hand, it might lend weight to a statement that might otherwise be considered unsupported (Latour and Woolgar, 1986, p. 80).

The role of certainty in academic texts has also been researched by linguists. A very influential line of research investigates the role of modal words and expressions that express an author's attitude towards what they are saying (e.g., could, may, possibility, clear, etc.). An often-used model to describe the use of linguistic devices to nuance claims in academic texts is that of hedging (Lakoff, 1973; Hyland, 1998). According to Lakoff (1973), who coined the term, hedges are "words whose job is to make things fuzzier or less fuzzy" (p. 471). Hyland (1998) built on this work to describe their role in the negotiation of academic knowledge. Hedges are a way "to balance objective information, subjective evaluation and interpersonal negotiation, and this can be a powerful persuasive factor in gaining acceptance for claims" (Hyland, 1998, p. 354). In general, hedges indicate either a lack of commitment to the truth value of a statement or a desire not to express that commitment unequivocally (Hyland, 2005). Hyland (2005) discusses three functions of hedges: to specify the actual state of uncertainty of a claim, to protect the writer against consequences of overstatements, and to appeal to the capability of readers to evaluate provisional statements. However, it is good to note that there is also an observed difference in the use of hedges between first and second language users. Second language students used a more limited range of hedges, offered stronger commitments, and showed greater difficulty with conveying a precise degree of certainty (Hyland and Milton, 1997).

Another relevant remark on the use of the model of hedging for assessing the conveyance of certainty of claims is made by Plappert (2019). He points out that the model overlooks the use of implicit meaning to convey (lack of) certainty. Plappert argues that the most typical type of claim made in the *Nature Genetics* articles he analyzed is a type four claim, without modalities (e.g., mutations in the gene encoding X *cause* disorder Y). He identifies only very few hedges, some of which do not even seem to convey uncertainty of the claim. However, he does identify multiple instances of implied causative relationships. In some instances, authors use a specific verb that signals possible causation (e.g., "associate"). Thereby, they seem to prefer unhedged claims that do not involve a speculation at the sentential level, but at the level of speaker meaning. With this observation, Plappert explicitly draws on Grice's theory of implicature (Grice, 1989). In short, this theory describes that there is a difference between what a speaker "says" (sentential level) and what they "implicate" (level of speaker meaning) (Grice, 1989). Plappert draws on Grice's conversational maxims to explain how scientists use implicature to make type one, two, or three statements at the level of speaker meaning, while making type four statements at the sentential level.

To summarize, these approaches focus on the structure, quality, and language of scientific arguments. They result in descriptions of writing practices without relating them to underlying epistemological notions that might be conveyed and constructed through them. In our framework, we combine and optimize these approaches from different fields to form phase 1 of epistemological notions analysis. By adding

five other phases, we proceed the analysis from descriptions of writing practices (argumentation analysis) to interpretations of what these practices can tell us about the epistemological notions they might convey and construct. However, before we explain these phases of the framework in detail, we first explain the concept of epistemological notions, relate it to views of NOS and epistemology, and point the reader's attention to some theoretical and philosophical underpinnings of the epistemological notions analysis framework.

2.2 Introducing epistemological notions underlying scientific language use and their relation with views of NOS

With our approach to analyzing students' scientific writing, we introduce a new concept in science educational research; epistemological notions underlying scientific language use. Epistemological notions, as we define them, are notions about the nature, origin and limits of knowing and knowledge conveyed and constructed through language use. To explain this definition, we explicate its parts and contrast it with the following related concepts: epistemology, views of NOS, and epistemological beliefs. An overview of the different concepts we discuss in this article and how they are related to the epistemological notions analysis framework can be found in Figure 1 and Table 2.

One aspect of the definition to explicate is "about the nature, origin and limits of knowing and knowledge"; the "epistemological" part of epistemological notions. A person's epistemology is their theory of knowledge; what counts as knowledge and how it comes about. Throughout history, various philosophers have formulated such theories. These theories differ in scope. Some have formulated ideas about how we can justify our beliefs in general (what we will call broad scope), some specifically write about how knowledge comes about in academic research contexts (medium scope), and others only describe how knowledge comes about in science (small scope). The broad scope encompasses all of a person's views about the nature of knowledge and the process of knowing in general and can include ideas about how they as a person determine whether their personal belief is justified (e.g., whether to trust authority or science textbooks as a reliable source). The medium and small scope only consider how we decide which beliefs are justified in academia. There is a difference between these two in which disciplines of academic research they include. A science specific epistemology (small scope) usually includes natural sciences and sometimes also social sciences. A general epistemology of academic research (medium scope) pertains to sciences and humanities, thus also includes disciplines in the humanities and interpretivist approaches to social sciences (Sundholm, 2014). This difference is especially relevant in our own research context since Dutch students learn about sciences and humanities as one concept, encompassing all academic research, with the word "wetenschap," also known from German as "Wissenschaft." In addition, students at Dutch universities are all trained to be academic researchers, "wetenschappers," since there are separate institutions for the applied sciences. So, Dutch universities are decidedly different from, for example, colleges or universities in the United States of America. We introduce the concept of epistemological notions, here, to apply it in research into language use in academic research (the medium scope of epistemology) or scientific research (small scope). The context of our own research is science (small scope), since the students we study and teach are students in the biomedical sciences.

In the second part of the concept, "notions," we chose to point toward the implicit and often incoherent nature of our construct of epistemological notions. They are latently present in writing or speech, and they are not necessarily part of a coherent epistemology. The difference with "ideas" or "views" about aspects of scientific knowing and knowledge is that epistemological notions in a text do not necessarily correspond with the writer's explicit beliefs about these aspects. Epistemological notions are conveyed and constructed through a certain way of writing (writing practices, Table 2). Why a writer has chosen this specific way of writing is probably the result of a combination of factors. These can include the writer's explicit ideas

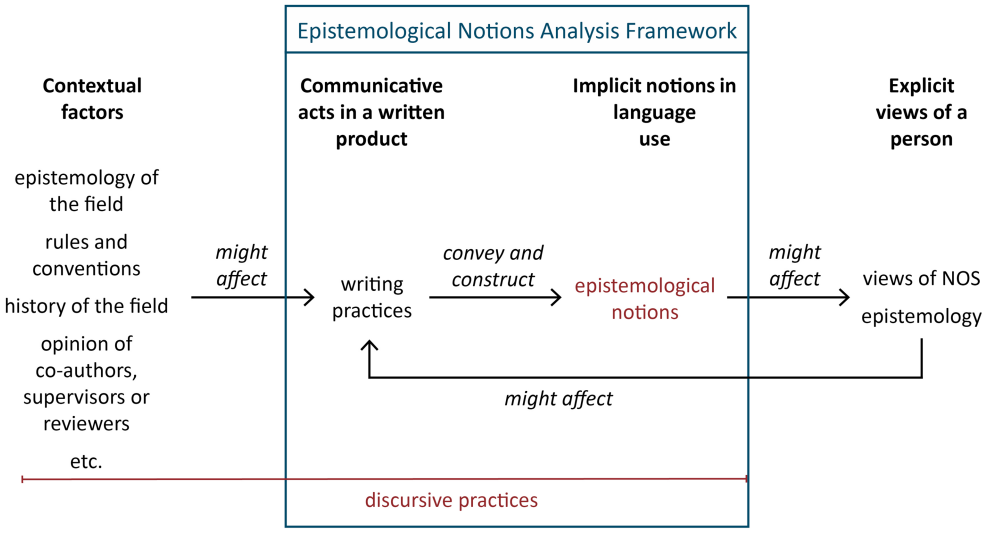

FIGURE 1
Relationships between concepts in the epistemological notions analysis framework.

TABLE 2 Explanation of concepts.

Concept	Explanation
Epistemological notions	Notions about the nature, origin and limits of knowing and knowledge conveyed and constructed through language use. They do not necessarily equate with the writer's or speaker's explicit views about knowledge and knowing.
Epistemology	A theory of knowledge; a theory about the nature, origin, and limits of knowing and knowledge.
Views of NOS	Views of the nature of science, detailing what a person believes about (the epistemological underpinnings of) scientific knowing, as a set of activities, and scientific knowledge, as a result of those activities (Lederman, 2007). Most often used for explicit, professed views. Most commonly describes only a specific subset of characteristics of science, called the consensus view of NOS (Abd-El-Khalick and Lederman, 2023).
Epistemological beliefs	A person's personal beliefs about knowledge and knowing (Hofer and Pintrich, 1997).
Epistemic	Relating to knowledge. "Epistemic level" relates to the degree of abstractness of knowledge claims, from grounded, low inference claims to progressively more general, theoretical claims (Kelly et al., 2007). "Epistemic status" relates to the degree of certainty awarded to knowledge claims.
Writing practice	A communicative act or a description thereof. Communicative acts can be words, phrases, sentences, text structure, argumentative moves, the act of combining any of those in a specific way (e.g., using two words interchangeably), and the inclusion of certain information or the omission of it.
Discursive practice	Description and interpretation of a communicative act in its context. It describes the communicative act itself and the interpretation of the researcher about the meaning of this act in its context.
Discourse	Language in use. It concerns how language is used to create meaning, for example, how it is used to say, do, and be things. Discourse analysis is the analysis of language in use, studying connections among and across sentences and studying how context gives meaning to language use and how language use gives significance to context (Gee, 2014).

about NOS and epistemology, but, for example, also the epistemological assumptions, history, and rules and conventions of their field of research and the instructions of co-authors or teachers (Figure 1). With our discourse analytic approach, we take this context into account in our interpretation of writing practices (resulting in discursive practices, Table 2). However, this remains the interpretation of the researcher doing the analysis. With this analysis we cannot make inferences about the writer's intentions or views. Therefore, we have explicitly included the following aspect in the definition of epistemological notions: "conveyed and constructed through language use."

It is also good to note the difference between epistemology and views of NOS. Not only is there a difference between their scope (knowledge in general or specifically scientific knowledge), there is also a difference in topics discussed. NOS is a construct created by science education researchers and educators with the specific purpose of identifying what, mainly K-12, students should know about the nature of scientific knowledge and knowing (Abd-El-Khalick and Lederman, 2023). The resulting frameworks are not epistemologies since they purposefully remain silent on certain aspects of epistemology (especially non-resolved controversies) and they are generally simplified, for pragmatic and didactic reasons. Reasons include time spent on NOS teaching and learning, and students' interests and readiness for learning (McComas, 2020). So, the construct of NOS addresses a specific set of aspects of epistemology. The NOS-framework that is used most often focuses on generally agreed-upon conclusions in a for students accessible manner. It is also good to note that it is still debated in the science education community which topics and views should be included in NOS teaching and learning (see Section 2.3). These debates also regard the question of whether or not to address non-resolved controversies in epistemology (Abd-El-Khalick and Lederman, 2023). Nonetheless, our concept of epistemological notions can relate to both epistemology and different conceptualizations of NOS, since it involves any notion relating to the nature, origin, and limits of knowledge and knowing.

A last concept related to epistemology to contrast epistemological notions and views of NOS with involves what is often called "epistemological beliefs." Epistemological beliefs are seen as a student's personal beliefs about the nature of knowledge and the process of knowing (Hofer and Pintrich, 1997). This often relates to their own knowledge, how they determine whether their own beliefs are justified (and thus can be seen as knowledge), and to their own learning, how a person learns and what their ability for learning is (Schommer, 1990). The concept of epistemological beliefs is mostly used within the field of educational psychology and has to do with students' cognitive development (Perry, 1968; Schommer, 1990; King and Kitchener, 1994; Kuhn, 1999). Views of NOS, on the other hand, describe (the epistemological underpinnings of) characteristics of scientific knowing, as a set of activities, and of scientific knowledge, as a result of those activities (Lederman, 2007). So, again, there is a difference in scope. Epistemological beliefs is a construct that uses epistemology in its broadest scope (beliefs about knowledge and knowing in general), while NOS is a construct that uses epistemology in the smallest scope (beliefs about scientific knowledge and knowing). Common elements in the construct of epistemological beliefs are beliefs about the certainty of knowledge (from fixed to more fluid), the simplicity of knowledge (from discrete, concrete and knowable to relative, contingent and contextual), the source of knowledge (from handed down by authority to derived from reason), and the justification of knowing (from right-or-wrong to a continuum of less or more justified beliefs) (Hofer and Pintrich, 1997). However, as Hofer and Pintrich (1997) concluded, there is no generally accepted definition of the construct of epistemological beliefs, and different authors include different additional elements in the construct. So, with some authors, there might be some overlap between epistemological beliefs and views of NOS. The epistemological notions analysis framework is mostly focused on beliefs about knowledge in academic research (medium scope

epistemology) or beliefs about knowledge in science (small scope epistemology), although it might also be applicable to define implicit notions in language use about personal knowledge, knowing, and learning (broad scope epistemology).

Although we contrast implicit notions and explicit views, we argue that there might be a relation between the epistemological notions underlying a writer's language use and their explicit ideas about epistemology. We believe, this relationship could have two directions. On the one hand, explicit ideas could engender epistemological notions in language use. In other words, what a writer thinks about how knowledge comes about can affect how they write about knowledge. On the other hand, epistemological notions in language use could affect explicit ideas. So, how a writer writes or what they read can affect what they think about how knowledge comes about. It is important to note that to assess these relationships, one needs to characterize both the epistemological notions in a written product or reading exercise (with the epistemological notions analysis framework) and the explicit views of the writer or reader, e.g., with a VNOS instrument (Lederman et al., 2002; Abd-El-Khalick et al., 2023), and explicitly study their relationship. We emphasize that the epistemological notions analysis framework is to be used only for that first part, characterizing epistemological notions in a written product or reading exercise. For research methods used to characterize a person's explicit NOS views, we refer the reader to the many works in science education literature addressing this. In the next section, we provide a short review of this literature.

2.3 A short review of views of NOS in science education literature

To study students' explicit NOS views, science education researchers have explicated characteristics of scientific knowledge and knowing which they believe are important for students to understand and act upon. Major NOS-frameworks are the so-called consensus framework (Abd-El-Khalick and Lederman, 2023) the Re-conceptualized family resemblance approach to NOS (RFN) (Erduran and Dagher, 2014; Kaya and Erduran, 2016; Cheung and Erduran, 2023), an approach focusing on questions rather than tenets (Clough, 2007, 2020), and the features of science approach (Matthews, 2012). Of these, the consensus framework is most used and referred to, followed by the RFN approach (Abd-El-Khalick and Lederman, 2023). Although the consensus framework is used most often, it has received many critiques over the years. The major critique is that it is oversimplified and too general to accurately reflect contemporary scientific practice in several sub-disciplines (Hodson and Wong, 2017).

For the consensus framework, science educators have formulated and substantiated the following list of consensus views, which forms the basis for NOS teaching and learning: scientific knowledge is empirical, inferential, creative, theory-driven, tentative, and socially negotiated. In addition, it includes the view that the existence of a single "Scientific Method" is a myth, that science is socially and culturally embedded, and that there is a difference between scientific theories and scientific laws (Abd-El-Khalick et al., 2008). However, it is good to note that this list of consensus views should not be treated as an exclusive, nor exhaustive list (Lederman, 2007). As we show with our use of the epistemological notions analysis framework, there are additional aspects of science and scientific knowledge about which people or texts can convey different views or notions (e.g., about the role of statistical outcome measures in scientific knowledge production).

The critique of the RFN approach to NOS is that there is not a single set of consensus views that fits with all scientific sub-disciplines and scientific enterprises. Therefore, it depicts science as a family of disciplines, each having its own characteristics as well as general features (Irzik and Nola, 2011). The key components of the RFN include the aims and values of science, methods and methodological rules, scientific knowledge, scientific practices as well as the social-institutional dimensions of science including the scientific ethos, professional activities, social certification and dissemination, social power structures, political power structures, financial systems, and social organizations and interactions (Erduran and Dagher, 2014).

It is important to realize that the list of consensus views is merely a short summary of more nuanced views of NOS and that these denote aspects that might not be apparent to our reader. We, therefore, refer the reader to other NOS work for an explication of these views (Lederman et al., 2002; Abd-El-Khalick, 2012; Abd-El-Khalick and Lederman, 2023). The presented key components of the RFN require more extensive explication as well. Such an explication can, for example, be found in the work of Erduran et al. (2019).

Furthermore, there are several other critiques directed at the consensus approach to NOS. The general critique is that it presents a too narrow focus (Allchin, 2017; Alsop and Gardner, 2017; Berkovitz, 2017; Hodson and Wong, 2017; Osborne, 2017; Simonneaux, 2017). Elements missing are the nature of scientific inquiry (understanding scientific practice) (Hodson and Wong, 2017), learning the language(s) of science (Alsop and Gardner, 2017; Hodson and Wong, 2017), controversies in epistemology and differences between subdisciplines (Hodson and Wong, 2017; Osborne, 2017; Simonneaux, 2017), a historical perspective for understanding (contemporary) science (Allchin, 2017; Berkovitz, 2017; Osborne, 2017), and a critical social justice view on science and NOS (Bazzul, 2020). Further, the focus of the consensus framework on demarcating science from pseudoscience with a static list of features is criticized. Instead, NOS-education should focus on the practical skill of knowing how to address NOS-issues as they emerge (Allchin, 2017), on different styles of reasoning in science (Osborne, 2017), or on students' own philosophical reflection on the foundations of sciences (Berkovitz, 2017). Simonneaux (2017) proposes to use two complementary approaches: a macro-level approach based on the consensus framework to teach what Latour describes as ready-made science, and a micro-level approach based on critical analysis of contemporary, controversial issues to teach about science-in-the-making.

For the current article it is good to note that the epistemological notions framework is theoretically flexible with respect to the approach to NOS. We believe, it can be used with a consensus view approach, an RFN approach to NOS and other more pluralistic approaches.

2.4 Theoretical and philosophical underpinnings

Although the epistemological notions framework is theoretically flexible with respect to the user's approach to NOS, there are three

theoretical assumptions underlying the framework that are essential to its use. First, we see language as constructing things in the world (e.g., Gee, 2014). Language is used to inform each other, but it is also used to do things and to be things (Gee, 2014). When one writes "The data were analyzed with SPSS," they are informing the reader about the program used for analysis. However, by writing it from the perspective of the data, the writer is also engaging in the practice of writing objectively about scientific procedures. In addition, because of their way of writing, the reader can identify the writer as an academic researcher. So, language is used to say, do, and be things. Therefore, it does not merely express things that already exist, it also creates things such as practices (writing objectively) and identities (researcher). In addition, language use does not only construct *ideas*, but it also constructs significance, practices, identities, relationships, politics (the distribution of social goods), connections, and sign systems and knowledge (Gee, 2014). This way of thinking about the constructive acts of language[1] is also important for using the epistemological notions analysis framework. Through discourse analysis one characterizes what communicative acts in their data might convey to and construct in the world. These descriptions and interpretations of writing practices are called discursive practices (Table 2). Through these constructive acts of language, the way we and others speak and write can affect our views of a topic. That is not to say that it always will, but it is good to be aware that it might change our views. This is especially relevant in the context of science education.

Second, and closely related, we see meaning as socially constructed (e.g., Latour and Woolgar, 1986). To provide meaningful characterizations of language use, we should use a critical approach in analyzing a writer's texts. A writer's language use cannot be characterized in isolation. We need to consider the discourses they are part of, and the ideas they might have about "proper" language use in scientific texts. In addition, other people than the writer can affect the final text, for example, informal and formal peer reviewers. Therefore, context is important in our analysis of language use. Furthermore, for students, we need to consider the fact that they are not just reporting scientific research but by doing so they are *learning* how to adequately report research as well. So, students not only might express the views they hold in the way they write about science, the way they write about science could also build their views. For example, when we teach students to write objectively, they could develop the view that a scientist's identity is unimportant and does not affect their science (relating to the theory-driven, creative, socially negotiated, and socially and culturally embedded NOS).

Third, in using the epistemological notions analysis framework, we see the subjectivity of the researcher using the framework as vital to qualitative analysis. A researcher's interpretations of epistemological notions underlying scientific language use are guided by the researcher's beliefs and feelings about the world, and their experience in it. In qualitative approaches, the researcher is the instrument for analysis (Nowell et al., 2017). We can only make sense of epistemological notions underlying scientific language use by being a person with epistemological and NOS views ourselves. Therefore, the epistemological notions analysis framework is not a step-by-step method that one follows to objectively characterize epistemological notions in scientific language use. The researcher is the instrument of analysis, and the framework is a tool to guide the researcher in systematic interpretation of a dataset to construct knowledge about the epistemological notions conveyed and constructed by the text. For any research instrument, it is important to examine how it works, what its underlying assumptions are, what its limitations are, and with which paradigms it does and does not fit. Since, here, the researcher is the instrument, careful and continuous self-examination, or reflexivity (Braun and Clarke, 2022), is key to using the epistemological notions analysis framework. Although the analysis framework is not a step-by-step method, we did implement the six phases of reflexive thematic analysis in the framework. Structuring qualitative data analysis by these phases can help a researcher in systematically and robustly exploring, interpreting, and reporting their data and analysis (Braun and Clarke, 2006, 2021). We elaborate on the use of researcher subjectivities as an analytical resource and the importance of continuous reflexivity in Section 4.2 of this article.

Because of these three theoretical assumptions, we see epistemological notions as situated notions that are brought about through discourse rather than as individual properties held by individuals. In other words, with this analysis framework, we focus on epistemological notions in writing products instead of NOS views held by a person (Table 2 and Figure 1). However, we do argue that student views can indeed be *formed by* their language use, actions, and behaviors, and that their views are context dependent (in contrast with language, actions, and behaviors as mere expressions of already formed conceptions).

In addition, it is good to realize that students generally do not have a coherent epistemology, nor do most scientists. So, when they behave in a certain way, that way may fit with a certain epistemology, but it does not mean that the student subscribes to that epistemology when you ask them about it or that they will always act according to that epistemology. Therefore, this analysis tells us something about a student's actions and how these actions can convey and construct notions relating to epistemology. This might be related to their explicit views of how knowledge comes about in science but does not necessarily have to be the case. Let alone that these epistemological notions fit with a coherent system of beliefs relating to epistemology or NOS.

2.5 Summary: the building blocks for the epistemological notions analysis framework

The epistemological notions analysis framework is based on descriptive text analysis approaches of the structure, quality and language of arguments (discussed in Section 2.1), and two interpretative analysis approaches, which are discourse analysis (Gee, 2014) and reflexive thematic analysis (Braun and Clarke, 2006, 2021). Together, these approaches are combined to characterize epistemological notions underlying scientific language use. This results in two types of analytic outputs. First, an overview of *what*

1 We chose to call Gee's (2014) "building tasks" in our text "constructive acts of language," because of the parallel with the constructivist paradigm underlying this framework and because "task" implies agency on the part of language, while it is people who construct things *with* language.

epistemological notions are conveyed and constructed in the text. Second, an overview of discursive practices, detailing *how* language use conveys and constructs epistemological notions. The researcher using the epistemological notions analysis framework characterizes what writing practices might say about the writer's ideas, presuppositions, identity, what they think is significant, how they treat the reader, and the ways of knowing they privilege (i.e., formulating discursive practices). From these discursive practices the researcher can characterize what epistemological notions might be conveyed or constructed through them. These interpretative steps take epistemological notions analysis a step further than previously published argument analysis studies in educational research (Section 2.1).

3 Development of the framework

3.1 Context of study

To develop the epistemological notions analysis framework, we have chosen to use student bachelor theses because these are culminating assessments which are used to assess students' knowledge and skills regarding scientific research and scientific writing. They are examples of how we have taught students to write and they regularly have a strong resemblance to the scientific articles they have read throughout their studies. Therefore, they can provide a rich source for educators to reflect on the epistemological notions implicitly conveyed and constructed at university and on how these notions might affect students' views of NOS and views relating to epistemology. We have chosen to use theses of biomedical students because this is our own disciplinary expertise.

Theses were written during an internship at the end of a 3-year bachelor program Biomedical sciences at Utrecht University in The Netherlands. They were accredited with 15 European Credits upon completion. Students worked in the lab for approximately 2 weeks and for 8 weeks they performed a literature study and wrote their thesis (minimum of 30–35 pages). These theses contain a literature study (minimum of 25 pages) and a separate research report in the form of a science journal article. Students were approximately 20–23 years old and had some course-based laboratory experience, but most had not participated in authentic research before. Most of the students followed only biomedical sciences or biology courses. Approximately 97% of the students continue with a Biomedical Research Master after their Bachelor program. Of those students, approximately 50% remain in research after their master's. So, biomedical sciences bachelor programs in the Netherlands are highly research oriented.

3.2 How we developed the framework

A total of nine theses were used for the development of the analysis framework. Five of these were written in 2017/2018 and 2018/2019 (archived theses) and four in 2021/2022 (newly obtained theses). We started out with the archived theses. For selection of archived theses for stages 1 and 2 of the development of the analysis framework, we used random sampling. For obtaining new theses for stage 3 of the development, we used purposive maximal variation sampling. We refer to our preregistrations (Pieterman-Bos et al., 2021, 2022) for details about the sampling and case selection strategy and to the Supplementary files for an elaborate description and justification of the development of the framework.

4 The epistemological notions analysis framework

4.1 A six-phase process to guide epistemological notions analysis

The epistemological notions analysis framework describes a process that guides the researcher using the framework in analysis of written scientific texts with the aim of characterizing discursive practices and the epistemological notions constructed by these discursive practices. The six phases of the process are inspired by the six phases of reflexive thematic analysis (Braun and Clarke, 2006, 2021). They are, however, different in content since they are focused on characterizing epistemological notions rather than the more general concept of "themes." So, reflexive thematic analysis is focused on identifying patterns of shared meaning in the way participants speak or write about a specific topic. Epistemological notions analysis, on the other hand, is focused on identifying patterns in writing practices that convey the same notion relating to the nature, origin and limits of knowing and knowledge. In addition, the framework builds on descriptive text analysis and discourse analysis approaches and thereby it puts stronger focus on the writing practices from which the epistemological notions are constructed.

We present the framework as an instruction guide for researchers who might engage in epistemological notions analysis in the future. We provide a guide through the six phases and offer examples to demonstrate how a researcher can use the framework for analysis of a scientific text. In the current article, the examples have a merely illustrative purpose. So, the interpretations presented here should be seen as a starting point to showcase this type of analysis. Further research, using this framework to analyze scientific texts should be carried out to substantiate which epistemological notions are conveyed and constructed through scientific language use.

The different phases of the framework are summarized in Figure 2. It is important to recognize that the framework is a tool that can guide the researcher in the process of analysis. It is not used as a step-by-step linear method that automatically leads to the emergence of epistemological notions. Like themes (Braun and Clarke, 2006, 2021), epistemological notions are active co-constructions on the part of the researcher, the written product, and context. In addition, using the epistemological notions analysis framework is a recursive process, during which a researcher moves back and forth between different phases. For example, while a researcher is writing about the analysis (phase 6), they will often go back to refining and defining the epistemological notions they characterized (phase 5), since writing often leads to further insights about the things one is writing about. They might also see reasons to characterize new discursive practices (phase 2) or epistemological notions (phase 3), which they then further develop and refine through phases 4 and 5. So, the framework should not be seen as a rigid, step-by-step method but as a tool for a systematic approach to data analysis.

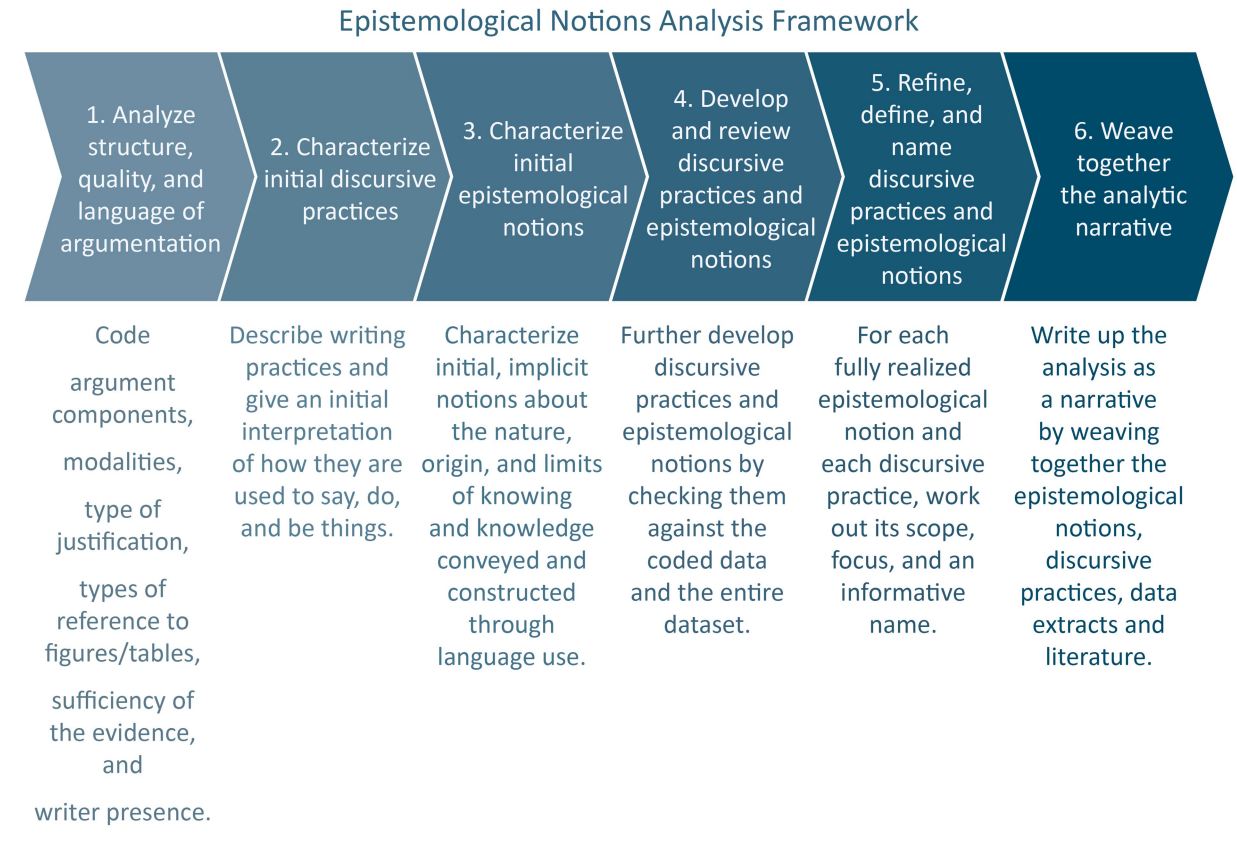

FIGURE 2
Phases of the epistemological notions analysis framework.

For each of the phases we offer examples. The data extracts from the theses we analyzed for the development of the framework often require specific content knowledge or elaborate introduction. So, for ease of reading, we have chosen to sometimes adapt data extracts (e.g., changing content-specific details to general terms) to make it more comprehensible or we have chosen to provide descriptions of data rather than the data extract itself. In addition, we use some fictitious examples to show the breadth of possible writing practices. Again, we acknowledge that further research is necessary to provide a full picture of which epistemological notions are conveyed and constructed through scientific language use. However, that is not the focus of the current article. As a note beforehand, the researcher can use the epistemological notions analysis framework both for characterizing epistemological notions of one writing product (one thesis) and for characterizing epistemological notions common in a type of writing product (a set of theses). These are different strategies that determine whether the researcher constructs patterns throughout one text or the entire dataset.

4.1.1 Phase 1: analyze structure, quality, and language of arguments

Before starting with analyzing the structure, quality, and language of arguments in phase 1, we advise researchers to take the time to familiarize themselves with the data. This involves repeated reading of the data. For more complex scientific texts, this also often involves reading secondary sources about the topic of the text to make sure you are familiar with the content.

After repeated, careful reading, the researcher proceeds to analyze the structure, quality, and language of the argument. For each argument aspect, the researcher labels parts of the text with codes that are fitting to that part. See Table 3 for an overview of all the argument aspects and their corresponding codes. It is good to note that since the codes relate to various aspects of the argument made, a single data extract can be labeled with multiple codes. There are two aims for this coding phase. The first aim is to reconstruct the argument made by the writer. To that end, the researcher identifies the different argument components: figures or tables, evidence sentences, justifications, qualifiers, and claims. The researcher can then reconstruct the argument structure, for example, by making a flowchart to visualize how the argument components relate to each other. The second aim of this phase is a more general aim of coding, that is, to organize the data into meaningful groups. These codes make it easier to navigate the data because they group similar writing practices. In the explication of coding for each of the argument aspects below, we will be concise about the argument aspects and codes that have been described in previous literature and we refer to Table 1 and the literature referenced there for details about these aspects and codes. We will focus our description on the newly defined aspects and codes (indicated with an asterisk in Table 3).

The researcher starts with labeling the argument components. We have slightly refined some of the descriptions of the argument components defined by Jiménez-Aleixandre and Federico-Agraso (2009) (Table 1) but refer to their work for a more elaborate description

TABLE 3 Description of adapted argument aspects and codes used in phase 1.

Argument aspect	Code	Description
Argument component	Figure or table	The tables and figures of the article that contain research data.
	Evidence	Sentences about the experimental results supporting the claims.
	Justification	Establishes the connection between a figure, table or evidence sentence and a claim (the equivalent of Toulmin's warrant). Justifications can be direct, being a simple justification in support of a claim, or subsequent, being both the endpoint for one argument (i.e., a claim) and a connection between evidence and claim in a second argument (i.e., a justification).
	Qualifier	A sentence modifying the degree of certainty of a claim or evidence sentence. Some qualifiers are deduced from data (e.g., confidence levels), others are deduced from previous research.
	Claim	A thesis the authors are seeking to demonstrate. Different levels of claims can be discerned: title claims made (often repeated) in section titles, intermediate claims made throughout the text, and end claims made in the conclusion section or abstract.
Modality	Hedge	Words/phrases used to decrease the certainty of a statement.
	Booster	Words/phrases used to increase the certainty of a statement.
	Implicature	Words used to speculate at the level of speaker meaning (what a writer implicates) while at the sentential level (what a writer writes) the statement has a higher level of certainty. Example: "Gene X is associated with disease Y." At the level of speaker meaning, it suggests that Y *might* be caused by A.
Type of justification used for a claim	Authority	Instances where a writer explicitly states a source of authority or lack thereof (e.g., teacher, class, book).
	Reasoning*	Justifications that refer to reasons based on a theoretical concept, or explanation of a theoretical concept, or explanation of a method.
	Empirical	Justifications citing some kind of empirical evidence or lack thereof (e.g., literature reference, data).
	Factual	Justifications written down as statement of fact without backing. Often repeating the original claim by using the exact same words, paraphrasing, or rephrasing.
Type of rhetorical reference to figures and tables	Inclusion	A figure or table is included in the explanation without reference to it in the text. When a claim or evidence sentence is obviously derived from a specific figure or table but not referred to in the text, label it with both inclusion and the other code that describes the type of analysis made with the data.
	Pointer	A non-descriptive reference to a figure or table in the text (e.g., "See graph 1"), without introduction, description, assertion, or interpretation.
	Introduction*	A sentence introducing a figure or table. For example, to indicate why data were included, to explain what experiment led to the figure, or to explain how to interpret this type of figure or graph.
	Description	A summary or other description of the figure or table with no suggestion of its relation to a claim. It does not go beyond what can be directly observed from the data inscription (without content knowledge).
	Assertion	A sentence about a figure or table in which the figure or table is asserted to show or prove a claim, without an explanation as to how it does so. It goes further than what can be directly observed from the table or figure but does not relate specific features of the data to the claim.
	Interpretation	A sentence that explicitly relates specific features of a table or figure to a claim. It goes further than what can be directly observed from the table or figure and requires an interpretation of its meaning in the context of the study.

(Continued)

10.3389/feduc.2024.1452111

TABLE 3 (Continued)

Argument aspect	Code	Description
Sufficiency of evidence and justification	Sufficient	The evidence presented is sufficient to support the claim made in its presented state (i.e., it may be a qualified or hedged claim which requires less evidence) and the claim is sufficiently justified.
	Partial or no evidence*	There is insufficient evidence presented to support the claim.
	Partially or unexplained evidence*	The writer presents results which they do not take into account in the claim.
	Partial or no consideration of counterevidence*	There is counterevidence for the claim presented, but this is not refuted or explicitly weighed against the positive evidence.
	Partially or unevidenced underlying claim*	The claim is supported by a previous claim that is not sufficiently evidenced.
	Epistemic status mismatch*	The claim is given a stronger epistemic status than the evidence can substantiate.
	Lack of statistical significance*	Statistical significance is not reached for the results, but they are still used to support a claim.
	Unclear*	The statement is unclear and therefore assessment of sufficiency of the evidence is not possible. For example, it is unclear where a demonstrative pronoun refers to.
Writer presence*	Procedure animate specific	Referring to a person or group when describing scientific procedures. For example, "we stained cells," "cells were stained by [ref]."
	Procedure animate common	Implying involvement of a person or group when describing scientific procedures. For example, "one can stain cells," "the intention was to stain cells," "to stain cells," "when staining cells."
	Procedure inanimate	Writing about scientific procedures from the perspective of a procedure or object without personal involvement. For example, "cell staining is done," "this process was repeated," "the experiment was done."
	Results animate specific	Referring to a person or group when describing results. For example, "we show," "increased infection was shown by [ref]," "[ref] showed."
	Results animate common	Implying involvement of a person or group when describing results. For example, "you can see that cells are increasingly infected," "it can be seen," "this is what would be expected," "expectations are," "it was hard to interpret."
	Results inanimate	Writing about results from the perspective of data, results, a figure or table, or an object. For example, "the data show that cells are increasingly infected," "Figure 8 shows," "the results show," "the assay shows," "cells show increasing infection," "cells are increasingly infected."

*Argument aspects and codes indicated with an asterisk are newly defined compared to the ones described in previous literature (Section 2.1).

10.3389/feduc.2024.1452111

of how to use them as codes. An argument component that does require some more attention is "qualifier." We recognize (at least) two types of qualifiers that moderate the degree of certainty of the presented claims. With the first type, the writer points to results of earlier research to moderate the degree of certainty about the presented results. An example could be "This association has not been found previously." With the second type, the writer points to a result of data analysis to moderate the degree of certainty of the claim made. An example of this could be "However, the confidence interval was very wide (95% CI [0.01–10.34])." Although these sentences draw on different sources, they are both moderating the degree of certainty the writer expresses about the presented claim. Therefore, they fall into the category of "qualifier" described by Toulmin (1958).

The next argument aspect that the researcher codes for is "modality." Modality has been extensively described by Hyland and Milton (1997), Hyland (1998, 2005), and Plappert (2019). So, we refer to their work and Table 3 for a description of this argument aspect. What is important to emphasize here is that not all words or phrases that could function as a hedge or booster do indeed function as such. An emblematic example of a word that can be used as a booster but is not always used as such is "significant." In common language it usually refers to importance, but in academic texts it usually refers to statistical significance, which is often used in a more neutral sense. An example of a word that could but not always does function as a hedge is "possible." For example, in the phrase "linear regression makes it possible to determine (…)," the word does not decrease the certainty of the statement, while the phrase "the possible clinical impact," it does decrease the certainty. Therefore, we again emphasize that context matters in this type of analysis. Another observation about the importance of context that we want to make is about the use of the word "should." There are cases where "should" is used instead of "must," which leaves room to not do or to not believe what is expressed in the sentence, i.e., a hedging function. In that sense, it is used as "it should be so, but we are not entirely sure." On the other hand, there are cases where "should" is used instead of "can," which implies that what is expressed in the sentence ought to be done and does not leave room for other interpretations, i.e., a boosting function. In that sense, it is used as "it should be so, there is no other option." Therefore, the researcher can label one word in one context as "hedge," in another as "booster," and in yet another they can decide to not label it at all.

Next, the researcher returns to the sentences and phrases that they labeled with "justification." For each of the justifications, they determine the type of justification according to the description provided in Table 3 and Sandoval and Millwood (2007). While developing the current framework, we also identified justifications in the theses for which we found it clear that they were based on literature. These would be empirical justifications. However, they did not have (clear) references to that literature. Therefore, we decided to code them with both empirical and factual reference. We made sense of these writing practices during later phases of the analysis.

Then, for each sentence about data that is presented in a figure or table, the researcher determines what type of reference is made to the figure or table. A type of reference that is newly included in the epistemological notions analysis framework compared to Sandoval and Millwood (2005) is "introduction." These types of reference to figures or tables are just that; they are introductions to it. For example, the writer indicates why the data were included, describes the experiment that led to the data, or explains how to interpret this type of graph or figure. After labeling all these sentences that relate to data presented in figures or table, the researcher determines for each figure and table the highest level of reference that is made to it in the text. The researcher does this for each claim that is supported with data from the figure or table. So, a single figure or table can be labeled multiple type of reference codes.

The next argument aspect that the writer analyzes is the "sufficiency of evidence and justification." For each claim, they determine whether it is sufficiently evidenced and justified by the writer. When they deem it to be insufficiently evidenced, the researcher then determines the source of this insufficiency. This is a new step compared to labeling the argument aspect as described by Jiménez-Aleixandre and Federico-Agraso (2009) under the label "coordination of evidence across epistemic levels." There can be different reasons to label a claim as insufficiently evidenced or justified that could be of interest for characterizing the possible underlying epistemological notions. The first possibility is that the writer presents no evidence or too little evidence to support the claim. Then, there can be cases in which there is evidence presented that is not or not fully explained. Therefore, it is unclear whether this evidence supports the claim or not. A slightly different version of this is when the writer presents evidence that clearly contradicts the other presented evidence, but they do not refute it nor explicitly weigh it against positive evidence. A fifth case is when the writer builds upon a previous claim to support the new claim, while the previous claim is not sufficiently evidenced or justified. A sixth possibility is that the writer gives a claim a stronger epistemic status than the evidence can substantiate, which we call an epistemic status mismatch. The seventh possibility is that the writer uses results for which no statistical significance is reached in support of their claim. The last possibility is that the claim is formulated in such a way that it is unclear what is actually claimed by the writer. This can, for example, be the case if the writer uses a demonstrative pronoun of which it is unclear what it refers to.

Lastly, the researcher turns toward the argument aspect of "writer presence." The researcher discerns three different types of writer presence and in their labeling, they also discern sentences about procedures and sentences about results, thus resulting in six possible codes. The first type of writer presence is the animate specific type. This is applied to sentences where the writer uses a person or group as the subject of the sentence or if a person or group is mentioned in such a way that they played a role in the interpretation. Examples of results sentences are "I/we/they/(s)he show(s) that cells are increasingly infected," "Increased infection was shown by me/[ref]," "[ref] showed increased infection," "this leaves us not yet satisfied." Examples of sentences about procedures are "I/we/they/(s)he stained cells," and "cells were stained by me/[ref]." The second type of writer presence is the animate common type. This is applied to sentences where the writer implies human involvement but does not specify which human(s) are involved. Often, it takes the form of "one/you can see," but we also apply it to sentences of the form "it can be seen that cells are increasingly infected," "what would be expected," "expectations are," and "it was hard to interpret." We see the same type of sentences about procedures. Examples are "one/you can stain cells, "the intention was to stain cells," "it is interesting to stain cells," and "adding compound X." The last type of writer presence is the inanimate type. Here, the writer uses an inanimate object, concept, or action as the subject of the sentence. In sentences about results the subject can be a research object (e.g., "cells were increasingly infected"), but it can also be data (e.g., "the data show")

or a figure or table (e.g., "the figure shows"). In sentences about procedures, the subject is often a research object (e.g., "cells were stained"), or a procedure ("cell staining was done"). When writers use this third type of writer presence, they remove themselves or other people from the text. The researcher searches the whole text to label sentences about results or procedures with these writer presence codes.

Key advice for the researcher for this phase is to not spend too much time on labeling of data extracts for which it proves to be difficult to determine the code for a certain argument aspect. The researcher will revisit these data extracts in the following phases, which might help them determine the code then. Another piece of advice is to document the reasons for specific choices well. Since these codes are the ingredients for the other analysis phases, the researcher will revisit them later during analysis. In light of new insights, they might not be able to reiterate their arguments for choosing a specific code. Therefore, it is a good practice to keep track of the thought process that went into deciding which code to apply.

4.1.2 Phase 2: characterize initial discursive practices

The intended outcome of phase 2 is a list of discursive practices. In the analysis of written language, discursive practices are descriptions and interpretations of writing practices in their context. They describe the writing practice itself and the interpretation of the researcher about the meaning of this practice in its context.

The researcher starts with the interpretation of the writing practices that they labeled in phase 1. It is good to note, that one type of writing practice can have multiple connotations. For this phase, it can be helpful to use the discourse analysis questions about the constructive acts of language formulated by Gee (2014) as a lens to guide interpretation. A reformulation of these discourse analysis questions in the context of scientific texts is available as Supplementary material accompanying the online article. Examples of such questions are: "How are justifications used to build or withhold credibility of used methodology?" "How are hedges and boosters used to make the research or results less or more significant and in what way?"

After interpreting the writing practices that relate to the argument aspects described in phase 1, the researcher proceeds to attend to other writing practices that do not directly relate to the argument aspects. The researcher starts to read the full text attentively again and when they encounter a writing practice that appears interesting, they label it with a first code describing the writing practice. They then go through the rest of the text to potentially identify other instances of the same or similar writing practices and label these as well. The researcher can decide to immediately write down an interpretation of the writing practice to formulate an initial discursive practice. They can also do this in a later phase and decide to first focus on labeling writing practices with descriptions of the writing practice. Then, they can look for patterns in these writing practices later to interpret their meaning in their context in terms of underlying epistemological notions. We advise the researcher to take the time to go through the text, with the aim of labeling interesting writing practices, multiple times.

After multiple unguided rounds of labeling, the researcher will do some rounds of guided labeling. Here, the researcher can use each of the constructive acts of language as a guide and they can use the list of aspects that we have found to be interesting to attend to Table 4. An example of a guiding discourse question for this part is: "How are specific phrases used to privilege (post-)positivist or constructivist epistemologies?" Another specific round of labeling that can be useful, is one in which the researcher pays specific attention to what is not written. An example is the detail in the descriptions of the methods used. Sometimes, when a writer leaves out information (e.g., justifications for choice of methods), this can be informative for the researcher as well. It is easy to overlook these writing practices and therefore, a specific round of coding for them can be helpful. Together, these rounds result in a (probably long) list of initial discursive practices with corresponding data extracts (one or more per discursive practice).

4.1.3 Phase 3: characterize initial epistemological notions

The intended outcome of phase 3 is a list of initial epistemological notions that may underlie the writing practices described in phase 2. Epistemological notions are notions about the nature, origin and limits of knowing and knowledge conveyed and constructed through language use. An example of an initial epistemological notion in the theses we analyzed is: "conclusions are not certain but should be certain to have value." This is a notion that might underlie the

TABLE 4 Examples of writing aspects used in phase 2.

Writing aspect	Examples
Description of results	The writer's use of interpretative words, how the writer describes possible debates in the fields, or the words the writer uses to describe variables.
Writing about methods	The level of detail provided about methods, to what degree the writer explains the methods used, or whether the writer justifies the choice of methods.
Writing about statistics	Whether the writer describes negative results, whether the writer reports measures of spread or confidence, or whether relevance is discussed in relation to statistical significance.
Uncertainty or inconclusiveness	Whether the writer gives tentative explanations (e.g., using hedges) about inconclusive results, or whether they attend to uncertainty when describing previous research.
Writing about published or future research	How the writer writes about conclusions of published research, for example, as statements of fact, or as their own interpretation of the published results and conclusions. Whether they call for future research to attend to a current lack of empirical evidence.
Adherence to standard writing practices	Whether the writer follows written and unwritten rules of the discipline in writing style or diverges from these standard writing practices and in what way.

writing practice of writing a conclusion with both a hedge (decreasing its certainty) and a booster (increasing the certainty). On the one hand, the text then seems to convey that the results are inconclusive and need further validation. On the other hand, it seems to convey that this tentativeness would invalidate their results. So, it conveys the epistemological notion that only impressive, conclusive research is valuable.

Again, it is important to approach this phase systematically, to give each piece of the text and each interesting aspect of epistemology enough attention. There are different ways to systematize this part of the analysis (working through the text from top to bottom, or starting with the first identified discursive practice, etc.), the important thing is to choose a system and to stick to it. For example the researcher could start at the top of the text with the first discursive practice they identified in phase 2. Then they could look at all the quotes coded with this discursive practice to look for patterns and to determine which epistemological notion might underlie this discursive practice. Then, they can proceed to the next discursive practice to characterize underlying (initial) epistemological notions for each discursive practice.

Both for characterizing discursive practices (phase 2) and epistemological notions (phase 3), the key is to look for patterns in data extracts to identify common features in the writing practices. Identifying these patterns is the first step that helps the researcher to construct underlying notions that tie several writing practices together. This step helps to transcend mere description of writing practices by focusing on interpretation. This part is where our analysis framework leans heavily on reflexive thematic analysis, and we recommend the researcher to read about the difference between mere descriptions and interpretations in terms of views in the reflexive thematic analysis literature (Braun and Clarke, 2006). In reflexive thematic analysis, the equivalents of descriptions and underlying ideas are "topic summaries" and "themes." However, it is good to note that in these early phases, the researcher will probably first identify a long list of narrow patterns (initial epistemological notions). During the subsequent phases, they will pay more attention to identifying broader patterns, applicable to more data extracts at a time.

4.1.4 Phase 4: develop and review discursive practices and epistemological notions

The aim of this phase is to further develop the initial discursive practices and epistemological notions, to focus on finding broader underlying ideas (patterns) and to determine that they tell a convincing story. First, the researcher checks the discursive practices and epistemological notions against the already coded data. Second, they check them against the entire text or entire dataset (depending on the single text strategy or multiple texts of the same type strategy chosen). This can result in the decision to revisit phase 2 and/or 3 to characterize new discursive practices or epistemological notions. Third, the researcher compares the different discursive practices they constructed to see whether they can combine discursive practices that capture broader patterns in the writing practices. The same is done for epistemological notions. When discursive practices or epistemological notions are combined, it is important to check all corresponding data extracts again to determine whether the newly constructed discursive practice or epistemological notion still fits with all of them. This process can also result in the decision to split discursive practices or epistemological notions because the researcher decides they constitute different interpretative patterns. Data within each discursive practice or epistemological notion should be coherent and meaningful, while there need to be clear distinctions between separate discursive practices and between separate epistemological notions.

We repeat the warning that Braun and Clarke (2006) give for the equivalent phase in reflexive thematic analysis: avoid staying in this phase for too long. It is easy to keep fine-tuning the lists of discursive practices and epistemological notions. However, when the refining does not add substantial insight to the analysis, the researcher should stop. As Braun and Clarke (2006) mention, the process is comparable to editing a text; at some point the additional hour of fine-tuning does not add remarkable impact to the text anymore. That is the moment to proceed to the next phase. In addition, in this phase the researcher will probably need to overcome the fear of losing information when they decide to discard a discursive practice or epistemological notion or decide to combine two of them. Keep in mind that the aim is to answer the research question and thus to reduce data to descriptions and interpretations that are relevant for that question. Loss of information is inevitable; indeed, it is one of the aims of analysis.

4.1.5 Phase 5: refine, define, and name discursive practices and epistemological notions

When the researcher is satisfied with the constructed discursive practices and epistemological notions, they move their focus toward refining the corresponding names. For each discursive practice and each epistemological notion, the aim is to write a sentence that clearly reflects its essence. For discursive practices, we recommend the structure of "description-which-interpretation." An example from our analysis of the theses is: "Being prescriptive about how to conduct research [description] which leaves no room for the reader's views on how to conduct research [interpretation 1] and which supports the idea that choices in research design are right or wrong [interpretation 2]." In this case, the students wrote their recommendations for future research in terms of what "should" be done. With this writing practice, they leave no room for the reader's view, because they assert their recommendations as if they are the only right way to proceed. This seems to convey the notion that there is a single "right" way to conduct research and that there are "wrong" ways, rather than the notion that there are multiple reasonable alternatives.

Another aspect of phase 5 is that the researcher determines the boundaries of each discursive practice and epistemological notion. On the one hand, it must fit with each data extract that the researcher applied it to. On the other hand, it must not be too broad so that it also fits with other data extracts. However, one data extract can contain multiple writing practices and therefore, it can lead to the researcher's construction of multiple discursive practices and multiple possible underlying epistemological notions. The names (or rather sentences) need to immediately give the reader a clear impression of what the discursive practice or epistemological notion is about.

In addition to the names, the researcher writes down a longer description of each discursive practice and each epistemological notion. These descriptions can include (curated) examples of writing

practices that support them. We have found that examples can often more clearly show the boundaries of a discursive practice or epistemological notion than mere descriptions. However, the researcher also needs to make sure that they explain what unites these examples, what constitutes the pattern that they established.

An example of the development of a discursive practice from our analysis of the theses is the following. We saw that students used words as "demonstrate," "fact," "known to be," and "shown." At first, we coded each of these instances as a separate discursive practice under the initial name "use of [insert word used] builds science is fact notion." Later, we saw the pattern in these practices, merged them, and replaced the specific words with "booster." In addition, we specified how the booster was used, what it gave rise to. The use of the boosters increased the perceived significance of results. Further, we realized that these writing practices might not only construct the notion that science produces facts, but that they could also construct the notion that science is objective. The writer seems to hide their interpretation and rhetoric (their personal involvement) behind objective language use. "The figure demonstrates conclusion X," rather than "the data support inference X," or "our interpretation of the data in the context of our research question is X." Therefore, we renamed the discursive practice "booster to increase significance of results builds objectivity notion and science is fact notion." During further refinement in phase 5, we applied some changes to increase clarity and because of new insights in the corresponding epistemological notions. This resulted in the current phrasing: "Using a booster which attributes value to impressiveness of research results in the process of knowledge creation and which conveys the notion that data speak for themselves."

An illustrative example of the development of an epistemological notion from our analysis is the following. We saw patterns in how students described their methods and initially characterized two epistemological notions that could underlie these patterns: "Scientific methods do not have to be justified," and "Research design choices are right or wrong." In a later stage of the analysis, we reformulated the first to create a positively formulated sentence "Scientific methods speak for themselves." Grouping epistemological notions to look for broader patterns made us realize that these notions are related because they both seem to imply that the choice of the research design and the choice of methods is straightforward because you can either choose the right or the wrong method. So, it led us to decide that these two epistemological notions describe the same idea that the choice of research design and methods speaks for itself. We refined its name to "Choice of research design and research methods speaks for itself" during phase 5.

4.1.6 Phase 6: weave together the analytic narrative

In phase 6, the researcher's aim is to create a narrative about the discursive practices and epistemological notions that they characterized and to support the narrative with data extracts and findings from previous literature. Using the data, they put forth their argument for their answer to the research question.

For the choice of data extracts, it usually works best to choose those that best show the essence of what the discursive practice or epistemological notion is about, the most illustrative ones. In addition, it is easier for the reader when they do not need a too elaborate introduction or explanation. Still, sometimes an example nearer to the boundary of the epistemological notion can help to explicate that boundary and to distinguish the example from another writing practice or to distinguish one epistemological notion from another. However, the focus should not be on the data extracts themselves or the researcher's descriptions of them, the focus is the argument that the researcher puts forth about their research question, for which they use data.

The reason we include the write-up of the analysis in the analysis framework, is because we have found that it is part of the recursive process. Often during our writing, we have further refined discursive practices or epistemological notions or have changed their names because of new analytic insights. Making the argument on paper can be a very helpful step for the data analysis itself.

An important note about write-up, is that the researcher should keep in mind that this type of analysis is interpretative and constructive. In reflexive thematic analysis, themes do not emerge from data (Braun and Clarke, 2006). Similarly, discursive practices and epistemological notions do not emerge from a text. They are not pre-existing entities that a researcher finds in the text, they are their own interpretations of the writing practices. This also bears consequences for the vocabulary that the researcher uses in their article. For example, "the underlying epistemological notion *is*" or "the discursive practice that emerged" do not fit with the constructivist paradigm underlying this type of qualitative analysis. Instead, the researcher can use wordings like "the epistemological notion that could underlie this practice," or "this practice may derive/result from the notion that," or "the discursive practice we developed/identified/formulated/constructed/defined." Using this different vocabulary might be difficult for researchers unfamiliar with constructivist approaches to data analysis. Therefore, it is prudent to check the article specifically for these phrases.

4.2 Criteria for conducting good epistemological notions analysis

To make the resulting analysis valuable and transferable for the research community, there are some important criteria for conducting good epistemological notions analysis. For an elaborate description of these criteria, we refer to the works of Braun and Clarke (2006, Table 2, 2021) in reflexive thematic analysis since most of their criteria are also relevant for using the epistemological notions analysis framework. The main quality criteria in the context of epistemological notions analysis are the following. First, the epistemological notions analysis framework cannot be used as a step-by-step method that needs to be followed from phase 1 to 6. It is not a protocol for data analysis that needs to be strictly followed to lead to a predefined result. The analysis framework is a guide, a set of principles and instructions that helps a researcher to explore and interpret data systematically. The resulting analysis will be a product of the input of the researcher (their skills and perspective), the data (the scientific text), and the research contexts (e.g., context of data creation and context of interpretation). Secondly, researcher subjectivity is employed rather than abolished because it makes for a more informed reader, development of deeper insights about the writer's epistemological notions, and more informed claims about generalities and transferability. Lastly, continuous reflexivity and prolonged engagement with data are essential to increase the

quality and depth of analysis and to enable external evaluation of these aspects.

5 Discussion

5.1 What does the epistemological notions analysis framework offer educational researchers?

The epistemological notions analysis framework can be used as a qualitative analysis method to characterize epistemological notions underlying scientific language use. Contrary to NOS views approaches, this approach can shed light on how language use can imply epistemological notions, which in turn could possibly affect students' views of NOS, when NOS is not explicitly discussed in class. Epistemological notions analysis results in two types of analytic output. First, an overview of discursive practices in the written product. Second, an overview of what epistemological notions are conveyed and constructed through them.

The epistemological notions analysis framework is based on previously used descriptive text analysis approaches of the structure, quality, and language of arguments, discourse analysis, and reflexive thematic analysis. The unique combination of these three approaches results in a systematic approach to proceed from mere description of writing practices to interpretation of their meaning in relation to epistemological notions. Most NOS research is focused on describing what explicit NOS views students assert or claim to support when NOS is the explicit topic of a lesson, questionnaire, or interview (Deng et al., 2011). To those approaches we add a research procedure that is focused on interpreting what epistemological notions might be conveyed to students when NOS is not explicitly addressed. Both approaches are valuable for NOS research. We discuss the merits of the epistemological notions approach. In short, the approach responds to calls for more attention to context in NOS research (Deng et al., 2011; Abd-El-Khalick and Lederman, 2023) and attention to learning the language(s) of science (Alsop and Gardner, 2017; Hodson and Wong, 2017). It focuses on understanding factors that could affect what views of NOS students develop during their formal education.

Descriptive explicit NOS views approaches are often focused on measuring[2] students' NOS views at one point in time or to measure them before and after an educational intervention (Abd-El-Khalick and Lederman, 2023). These approaches are valuable to assess students' current NOS views or the impact of educational interventions on explicit views. However, from these studies, it is often difficult to draw conclusions about the learning processes that affect building and revision of students' NOS views. An interpretative approach, attending to context, can increase our understanding of what epistemological notions might be conveyed to students when NOS is not explicitly addressed. We have used the epistemological notions analysis framework to characterize epistemological notions conveyed and constructed in bachelor theses (publication in progress). We have also interviewed the students about their explicit NOS views and asked them to relate these to their own writing practices and scientific writing practices in general. A preliminary finding is that students sometimes have trouble with relating their explicit NOS views to the way they are taught to write. So, epistemological notions analysis can help researchers understand difficulties that students experience with translating professed NOS views to their writing. For example, students learn that scientific theories are formed based on multiple studies and that they are not substantiated by a single study. At the same time, students are instructed to add something new and substantial with their work. As a result, students could develop the idea that a single study should be conclusive to result in new and substantial knowledge, which contradicts the tentative (yet durable) nature of science. So, the epistemological notions analysis framework can also be used to characterize how the writing practices we teach could affect the NOS views students develop during their formal education. Although we remind the reader that this requires assessment of students' explicit views of NOS (e.g., using a VNOS instrument) as well.

By teaching students to adhere to certain rules and conventions in article writing, we also convey the epistemological notions underlying these discursive practices. Using the epistemological notions analysis framework sheds light on these epistemological notions-constructing discursive practices, and with that, it opens up a yet unexplored focus area for NOS research. For experienced scientists, community standards for writing about science might be more dissociated from their explicit NOS views. For example, when scientists write "the difference between A and B was statistically significant ($p = 0.03$)," they probably remain aware of the fact they themselves have set the alpha to 0.05 and that this bears consequences for type I and type II errors. Students, on the other hand, are more likely to see the statistical significance as a fully objective measure which has nothing to do with researcher choice in research design since they are still learning to grasp the concept of statistical significance. So, what we teach students about scientific discourse could affect their NOS views. Therefore, the epistemological notions analysis framework can be used by educational researchers to characterize epistemological notions that might be implicitly taught through writing practices. These insights can then be used to improve explicit-reflective NOS teaching and learning.

This approach, focused on language-in-use in science, can also offer a way to enrich consensus views of NOS with examples from the contemporary practice of science, as is argued for by Hodson and Wong (2017). By analyzing language use in science together with students, students could learn to be part of this scientific discourse and learn to meaningfully contribute to the language games that give meaning to science and possibly demarcate science (Alsop and Gardner, 2017). In such a lesson focused on scientific discourse, students could also learn to recognize how science is politically entangled. A discourse critical approach to scientific language could shed light on how science is socially, culturally, politically, historically, and materially situated, how it is entangled in forms of governance and power, and how it can contribute to social justice (or not) (Bazzul, 2020).

5.2 Some remarks on applicability of the framework

We have developed the epistemological notions analysis framework using university bachelor theses. It is good to note that these writing products are not only a product of the student,

2 The difference between approaches focused on measurement and those focused on understanding, which we discuss in this paragraph, is also discussed by Gough and Madill (2012).

their background, experience, and views. They are also a product of an internship, and therefore of the training and feedback the student's supervisor has given them and of the instructions given within the course. Therefore, the researcher using the framework should take these contextual factors into account in their analysis. The epistemological notions analysis could for example be complemented by interviews with the students. These could serve to map their experiences with supervision during the internship. In addition, interviews can explore students' professed NOS views to compare them with the epistemological notions conveyed through their writing. Lastly, specific writing practices can be discussed to ask students about their reasons for writing in a specific manner.

Although we have developed the framework using theses, we believe that it has broader applicability. Other writing products, where students use scientific data to support claims, can also be used for epistemological notions analysis. Many of the descriptive text analysis approaches used in phase 1 of the framework have also been used on different argumentation assignments. We advise researchers to consider which of these argument aspects (Table 3) are applicable to the assignment they want to analyze and to consider the contextual factors in play with that assignment.

In addition, the framework could be used to characterize the epistemological notions conveyed and constructed through science teachers' language use. The writing practices and NOS views of teachers and thesis supervisors could be an important factor affecting the NOS views that their students develop. As is also argued by Alsop and Gardner (2017), we need to pay closer attention to the ways in which our educational practices demarcate what gets to count as rigorous science and what does not. The results of an analysis of epistemological notions underlying science teachers' language use could be used to improve science teacher education programs and professional development programs for university teachers.

Lastly, we believe that it can also be valuable to use the framework for characterizing the epistemological notions conveyed and constructed through scientists' language use in published research articles. An analysis of the epistemological notions conveyed and constructed through scientists' language use could shed light on notions that are implicitly propagated through research articles or popular science writing. These might affect students' NOS views but could also affect NOS views of the general public when they read about science. Therefore, epistemological notions analysis might be an interesting approach for people interested in science communication, public engagement, or trust in science.

To conclude, the epistemological notions analysis framework can be used to characterize epistemological notions underlying scientific language use. We encourage the reader to use the framework and the outlined quality criteria along with more explicit approaches focused on professed NOS views. A combination of both approaches can help to understand how professed, explicit views relate to scientific (writing) practice.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: The data presented in this article are not readily available because Utrecht University has chosen to not make bachelor theses publicly available, and participants of this study did not give permission to do so. Requests to access these datasets should be directed to Marc van Mil, m.h.w.vanmil@umcutrecht.nl.

Ethics statement

The studies involving humans were approved by Ethics Review Board of the Faculty of Social and Behavioral Sciences Utrecht University, Utrecht, The Netherlands (no. 22-0204). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

AP-B: Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. CR: Conceptualization, Writing – review & editing. RS: Supervision, Writing – review & editing. MM: Conceptualization, Supervision, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

We thank Professor Fouad Abd-El-Khalick and our colleagues of the Life Sciences Educational Research program for their valuable feedback on several versions of the framework.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2024.1452111/full#supplementary-material

References

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. eds. B. J. Fraser, K. G. Tobin and C. J. McRobbie (Dordrecht: Springer), 1041–1060.

Abd-El-Khalick, F., and Lederman, N. G. (2023). "Research on teaching, learning, and assessment of nature of science" in Handbook of research on science education. eds. N. G. Lederman, D. L. Zeidler and J. S. Lederman (New York: Routledge), 850–898.

Abd-El-Khalick, F., Summers, R., Brunner, J. L., Belarmino, J., and Myers, J. (2023). Development of VAScoR: a rubric to qualify and score responses to the views of nature of science (VNOS) questionnaire. *J. Res. Sci. Teach.* 61, 1641–1688. doi: 10.1002/tea.21916

Abd-El-Khalick, F., Waters, M., and Le, A. P. (2008). Representations of nature of science in high school chemistry textbooks over the past four decades. *J. Res. Sci. Teach.* 45, 835–855. doi: 10.1002/tea.20226

Allchin, D. (2017). Beyond the consensus view: whole science. *Can. J. Sci. Math. Technol. Educ.* 17, 18–26. doi: 10.1080/14926156.2016.1271921

Alsop, S., and Gardner, S. (2017). Opening the black box of NOS: or knowing how to go on with science education, Wittgenstein, and STS in a precarious world. *Can. J. Sci. Math. Technol. Educ.* 17, 27–36. doi: 10.1080/14926156.2016.1271924

Bazzul, J. (2020). "Political entanglement and the changing nature of science" in Nature of science for social justice. eds. H. A. Yacoubian and L. Hansson (Cham: Springer), 79–95.

Berkovitz, J. (2017). Some reflections on "going beyond the consensus view" of the nature of science in K-12 science education. *Can. J. Sci. Math. Technol. Educ.* 17, 37–45. doi: 10.1080/14926156.2016.1271927

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Braun, V., and Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qual. Res. Psychol.* 18, 328–352. doi: 10.1080/14780887.2020.1769238

Braun, V., and Clarke, V. (2022). Conceptual and design thinking for thematic analysis. *Qual. Psychol.* 9, 3–26. doi: 10.1037/qup0000196

Cheung, K. K. C., and Erduran, S. (2023). A systematic review of research on family resemblance approach to nature of science in science education. *Sci. Educ.* 32, 1–37. doi: 10.1007/s11191-022-00379-3

Clough, M. P. (2007). Teaching the nature of science to secondary and post-secondary students: questions rather than tenets. *Pantaneto Forum* 25, 31–40.

Clough, M. P. (2020). "Framing and teaching nature of science as questions" in Nature of science in science instruction. ed. W. F. McComas (Cham: Springer), 271–282.

Cobo, C., Abril, A. M., and Romero-Ariza, M. (2022). Effectiveness of a contextualised and integrated approach to improving and retaining preservice teachers' views of the nature of science. *Int. J. Sci. Educ.* 44, 2783–2803. doi: 10.1080/09500693.2022.2151326

Deng, F., Chen, D. T., Tsai, C. C., and Chai, C. S. (2011). Students' views of the nature of science: a critical review of research. *Sci. Educ.* 95, 961–999. doi: 10.1002/sce.20460

Erduran, S., and Dagher, Z. R. (2014). Reconceptualizing the nature of science for science education: Scientific knowledge, practices and other family categories. Dordrecht: Springer.

Erduran, S., Dagher, Z. R., and McDonald, C. V. (2019). Contributions of the family resemblance approach to nature of science in science education: a review of emergent Research and Development. *Sci Educ* 28, 311–328. doi: 10.1007/s11191-019-00052-2

Gee, J. P. (2014). An introduction to discourse analysis: Theory and method. *4th* Edn. New York: Routledge.

Gough, B., and Madill, A. (2012). Subjectivity in psychological science: from problem to prospect. *Psychol. Methods* 17, 374–384. doi: 10.1037/a0029313

Grice, P. (1989). Studies in the way of words. London: Harvard University Press.

Hodson, D., and Wong, S. L. (2017). Going beyond the consensus view: broadening and enriching the scope of NOS-oriented curricula. *Can. J. Sci. Math. Technol. Educ.* 17, 3–17. doi: 10.1080/14926156.2016.1271919

Hofer, B. K., and Pintrich, P. R. (1997). The development of epistemological theories: beliefs about knowledge and knowing and their relation to learning. *Rev. Educ. Res.* 67, 88–140. doi: 10.2307/1170620

Hyland, K. (1998). Boosting, hedging and the negotiation of academic knowledge. *Text Talk* 18, 349–382. doi: 10.1515/text.1.1998.18.3.349

Hyland, K. (2005). Prudence, precision, and politeness: hedges in academic writing. *Quaderns Filologia Estudis Lingüíst.* 10, 99–112.

Hyland, K., and Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *J. Second. Lang. Writ.* 6, 183–205. doi: 10.1016/S1060-3743(97)90033-3

Irzik, G., and Nola, R. (2011). A family resemblance approach to the nature of science for science education. *Sci. Educ.* 20, 591–607. doi: 10.1007/s11191-010-9293-4

Jiménez-Aleixandre, M. P., and Federico-Agraso, M. (2009). Justification and persuasion about cloning: arguments in Hwang's paper and journalistic reported versions. *Res. Sci. Educ.* 39, 331–347. doi: 10.1007/s11165-008-9113-x

Kaya, E., and Erduran, S. (2016). From FRA to RFN, or how the family resemblance approach can be transformed for Science curriculum analysis on nature of Science. *Sci. Educ.* 25, 1115–1133. doi: 10.1007/s11191-016-9861-3

Kelly, G. J., and Chen, C. (1999). The sound of music: constructing science as sociocultural practices through oral and written discourse. *J. Res. Sci. Teach.* 36, 883–915. doi: 10.1002/(SICI)1098-2736(199910)36:8<883::AID-TEA1>3.0.CO;2-I

Kelly, G. J., Druker, S., and Chen, C. (1998). Students' reasoning about electricity: combining performance assessments with argumentation analysis. *Int. J. Sci. Educ.* 20, 849–871. doi: 10.1080/0950069980200707

Kelly, G. J., Regev, J., and Prothero, W. (2007). "Analysis of lines of reasoning in written argumentation" in Argumentation in science education. Science & Technology Education Library. eds. S. Erduran and M. P. Jiménez-Aleixandre (Dordrechts: Springer), 137–157.

Kelly, G. J., and Takao, A. (2002). Epistemic levels in argument: an analysis of university oceanography students' use of evidence in writing. *Sci. Educ.* 86, 314–342. doi: 10.1002/sce.10024

King, P., and Kitchener, K. (1994). Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults. San Francisco: Jossey-Bass.

Kuhn, T. S. (1970). The structure of scientific revolutions. Chicago: The University of Chicago Press.

Kuhn, T. S. (1982). "Commensurability, Comparability, Communicability," in *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, The University of Chicago Press, pp. 669–688.

Kuhn, D. (1999). A developmental model of critical thinking. *Educ. Res.* 28, 16–46. doi: 10.3102/0013189X028002016

Lakoff, G. (1973). Hedges: a study in meaning criteria and the logic of fuzzy concepts. *J. Philos. Log.* 2, 458–508. doi: 10.1007/BF00262952

Latour, B., and Woolgar, S. (1986). Laboratory life: The social construction of scientific facts. *2d* Edn. Princeton, New Jersey: Princeton University Press.

Lederman, N. G. (1992). Students' and teachers' conceptions of the nature of science: a review of the research. *J. Res. Sci. Teach.* 29, 331–359. doi: 10.1002/tea.3660290404

Lederman, N. G. (2007). "Nature of science: past, present, and future" in Handbook of research on science education. eds. S. K. Abell and N. G. Lederman (Mahwah, NJ: Lawrence Erlbaum), 831–879.

Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., and Schwartz, R. S. (2002). Views of nature of science questionnaire: toward valid and meaningful assessment of learners conceptions of nature of science. *J. Res. Sci. Teach.* 39, 497–521. doi: 10.1002/tea.10034

Matthews, M. R. (2012). "Changing the focus: from nature of science (NOS) to features of science (FOS)" in Advances in nature of science research: Concepts and methodologies. ed. M. Khine (Dordrecht: Springer), 3–26.

McComas, W. F. (2020). "Considering a consensus view of nature of science content for school science purposes" in Nature of science in science instruction. Science: Philosophy, history and education. . Ed. McComas, W. F. (Cham: Springer Nature), 23–34.

Metin Peten, D. (2022). Influence of the argument-driven inquiry with explicit-reflective nature of scientific inquiry intervention on pre-service science teachers' understandings about the nature of scientific inquiry. *Int. J. Sci. Math. Educ.* 20, 921–941. doi: 10.1007/s10763-021-10197-8

Nowell, L. S., Norris, J. M., White, D. E., and Moules, N. J. (2017). Thematic analysis: striving to meet the trustworthiness criteria. *Int J Qual Methods* 16, 1–13. doi: 10.1177/1609406917733847

Osborne, J. (2017). Going beyond the consensus view: a response. *Can. J. Sci. Math. Technol. Educ.* 1, 53–57. doi: 10.1080/14926156.2016.1271920

Perry, W. G. (1968) Patterns of development in thought and values of students in a Liberal arts college: A validation of a scheme. Available at: https://files.eric.ed.gov/fulltext/ED024315.pdf (Accessed December 6, 2023).

Pieterman-Bos, A., van Mil, M. H. W., and van de Schoot, R. (2021). Qualitative text analysis method for characterizing personal epistemology in student research articles [preregistration]. *Open Sci. Framew.* doi: 10.17605/OSF.IO/SC6YU

Pieterman-Bos, A., van Mil, M. H. W., and van de Schoot, R. (2022). Characterization of biomedical students' views of the nature of science as enacted in their own research practice [preregistration]. *Open Sci. Framew.* doi: 10.17605/OSF.IO/S8GRD

Plappert, G. (2019). Not hedging but implying: identifying epistemic implicature through a corpus-driven approach to scientific discourse. *J. Pragmat.* 139, 163–174. doi: 10.1016/j.pragma.2018.09.001

Sandoval, W. A., and Millwood, K. A. (2005). The quality of students' use of evidence in written scientific explanations. *Cogn. Instr.* 23, 23–55. doi: 10.1207/s1532690xci2301_2

Sandoval, W. A., and Millwood, K. A. (2007). "What can argumentation tell us about epistemology?" in Argumentation in science education. Perspectives from classroom-based research. eds. S. Erduran and M. P. Jiménez-Aleixandre (Dordrecht: Springer), 71–88.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *J. Educ. Psychol.* 82, 498–504. doi: 10.1037/0022-0663.82.3.498

Simonneaux, L. (2017). Au-delà de la polémique, compléter l'approche macro consensuelle de la NOS avec l'approche micro de la recherche en train de se faire. *Can. J. Sci. Math. Technol. Educ.* 17, 58–65. doi: 10.1080/14926156.2016.1271922

Sundholm, G. (2014). "The vocabulary of epistemology, with observations on some surprising shortcomings of the English language" in Mind, values, and metaphysics: Philosophical essays in honor of Kevin mulligan. ed. A. Reboul (Cham: Springer International Publishing), 203–208.

Takao, A. U., and Kelly, G. J. (2003). Assessment of evidence in university students' scientific writing. *Sci. & Educ.* 12, 341–363. doi: 10.1023/A:1024450509847

Toulmin, S. E. (1958). The uses of argument. Cambridge: Cambridge University Press.