

Framework for the natures of negativity in introductory physics

Suzanne White Brahmia^{1,*}, Alexis Olsho,¹ Trevor I. Smith², and Andrew Boudreax³

¹Department of Physics, University of Washington, Box 351560, Seattle, Washington 98195-1560, USA

²Department of Physics & Astronomy and Department of STEAM Education, Rowan University,

201 Mullica Hill Road, Glassboro, New Jersey 08028, USA

³Department of Physics & Astronomy, Western Washington University,

516 High Street, Bellingham, Washington 98225, USA



(Received 8 March 2019; accepted 17 December 2019; published 22 April 2020)

Mathematical reasoning skills are a desired outcome of many introductory physics courses, particularly calculus-based physics courses. Novices can struggle to understand the many roles signed numbers play in physics contexts, and recent evidence shows that unresolved struggle can carry over to subsequent physics courses. Positive and negative quantities are ubiquitous in physics, and the sign carries important and varied meanings. The mathematics education research literature documents the cognitive challenge of conceptualizing negative numbers as mathematical objects—both for experts, historically, and for novices as they learn. We contribute to the small but growing body of research in physics contexts that examines student reasoning about signed quantities and reasoning about the use and interpretation of signs in mathematical models. In this paper we present a framework for categorizing various meanings and interpretations of the negative sign in physics contexts, inspired by established work in algebraic contexts from the mathematics education research community. Such a framework can support innovation that can catalyze deeper mathematical conceptualizations of signed quantities in the introductory courses and beyond.

DOI: 10.1103/PhysRevPhysEducRes.16.010120

I. INTRODUCTION

A. Motivation

Experts in physics translate fluidly between different representations of phenomena. To an expert, a physics equation “tells the story” of an interaction or process. For example, when reading the equation $x(t) = +40 \text{ m} + (-5 \text{ m/s})t + \frac{1}{2}(-9.8 \text{ m/s}^2)t^2$, an expert may quickly construct a mental story of the covariation of position and time of a projectile that starts 40 m above the ground and is launched with a speed of 5 m/s vertically downward. Part of the challenge of learning physics is developing the ability to decode symbolic representations in this manner.

In these translation processes, experts can readily attribute specific meanings to positive and negative signs. In the example above, the positive sign in front of the term 40 m indicates that the projectile starts at a position that is in the positive direction from the origin, *upward* in this case because the gravitational acceleration is always downward and happens to be negative in this expression. The positive

sign after the 40 m term indicates that the following term, $(-5 \text{ m/s})t$, represents an additive change in the projectile’s position—a one-dimensional vector quantity that has the initial value 40 m and is in the positive direction. The sign in front of 5 m/s indicates that the projectile is launched *downward*.

Other examples of the fluid interpretation of signs abound in introductory physics:

- In the equation $\vec{F}_{12} = -\vec{F}_{21}$, the negative sign signals that the force exerted by object 2 on object 1 is in the exact opposite direction as the force exerted by 1 on 2.
- In the expression $0 - (-5 \mu\text{C})$, the first negative sign indicates that a quantity of electric charge is being removed from an electrically neutral object, while the second negative sign indicates which of the two different types of electric charge is being removed.
- In Faraday’s law, $\mathcal{E} = -d\Phi_B/dt$, the negative sign reminds the expert that the voltage induced by a changing magnetic flux acts to oppose (rather than reinforce) the change that created it.

In all of these cases, experts generally decode specific meanings of the sign quickly and effortlessly, perhaps in most cases without conscious awareness of the decoding process itself. Novices may need to spend considerable conscious effort interpreting the sign, or may fail altogether to successfully interpret it. Pitfalls likely to challenge the novice might include a tendency to overgeneralize a particular interpretation, or a lack of awareness that two nearby signs in an equation may have completely different

*brahmia@uw.edu

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](#). Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

physical interpretations. The challenge for introductory physics teaching may be compounded if the already difficult cognitive task of interpreting signs goes unaddressed when instructors themselves are not consciously aware of the mismatch between their own ability and the much lower skill level of novices.

An additional layer of complexity is introduced by the implicit nature of the sign of many symbolically represented quantities. For example, the expression ΔU_g is a stand-in for a generalized number of Joules, which could be positive or negative. The sign in this case would indicate whether the potential energy of some system increased or decreased during a particular process. U_g is also a stand in for a generalized number—but in contrast to ΔU_g , the sign of U_g tells an expert whether the energy of a system in a particular configuration is larger or smaller than the energy associated with a pre-established reference configuration.

In a fast-paced introductory physics course, this nuanced interpretation of sign may fall by the wayside, as instructors attend to a host of other, perhaps more obvious challenges. Unfortunately, student difficulties with decoding sign, if unaddressed, may not spontaneously resolve. In such cases, difficulties with signs could contribute to serious obstacles in the development of overall quantitative reasoning and analysis skills so highly prized by physicists. In this article, we describe an effort to systematically parse and document the various meanings of sign in introductory physics contexts. We hope that a taxonomy of this type can support instructor efforts to foster their students' ability to translate between representations, and can support the efforts of physics education researchers to more fully understand the nature of student reasoning about signs and signed quantities.

B. Background research

Negative pure numbers represent a more cognitively difficult mathematical object than positive pure numbers do for precollege mathematics students [1]. Mathematics education researchers have isolated a variety of “natures of negativity” fundamental to algebraic reasoning in the context of high school algebra—the many meanings of the negative sign that must be distinguished and articulated for students to develop understanding [2,3]. These various meanings of the negative sign form the foundation for scientific quantification, where the mathematical properties of negative numbers are well suited to represent natural processes and quantities. Physics education researchers report that a majority of students enrolled in a calculus-based physics course struggled to make meaning of positive and negative quantities in spite of completing Calculus I and more advanced courses in mathematics [4,5]. Developing “flexibility” with negative numbers (the recognition and correct interpretation of the multiple meanings of the negative sign) is a known challenge in mathematics education, and there is mounting evidence that reasoning

about negative quantity poses a significant hurdle for physics students at the introductory level and beyond [6–10]. Few published studies have focused on the use of the negative sign, i.e., *negativity*, specifically in the context of the mathematics used in physics courses. Studies conducted in the context of upper-division physics courses reveal robust student difficulties associated with interpretation of the vector nature of acceleration and its representation in Newton's second law, and with contexts in electricity and magnetism such as Coulomb's law, in which there are often multiple negative signs, each with a separate meaning [6,8].

Brahmia and Boudreux conducted studies to probe student difficulties with negative quantities at the introductory level. They constructed physics assessment items based on the natures of negativity from mathematics education research [11] and administered them to introductory physics students in an introductory sequence of courses [4,5,12]. They report that students struggle to reason about signed quantity in the contexts of negativity typically found in the introductory curriculum (e.g., negative work, negative direction of acceleration or electric field in one dimension), and they concluded that science contexts may overwhelm some students' conceptual facility with negativity. Olsho, White Brahmia, Smith, and Boudreux studied student understanding of the negative sign when used in the context of electric charge. On a multiple-response assessment item, students that correctly interpreted the meaning of a negative net charge were more likely to also correctly identify electrons as the charge carriers in the given situation—that is, correct interpretation of the negative sign in the context of negative net charge may be associated with a more complete understanding of the physical situation [13]. These studies reveal that signed quantities in introductory physics present cognitive difficulties for students that many do not reconcile even after completing the introductory sequence. These difficulties might then carry over into upper-division course work. Further, misinterpretation of a negative sign may be associated with a less complete understanding of a given physics context.

C. Contribution to the literature

The current study advances this body of research by introducing a framework for categorizing the natures of negativity in introductory physics (NoNIP), analogous to the natures of negativity developed in the context of algebra [11]. While we recognize that students struggle with signed quantities more generally, we choose to focus on negatively signed quantities in this work because they are the only signed quantities for which the sign is always explicitly included (e.g., a velocity in one dimension of 3 m/s is typically assumed to be in the positive direction). The intention is to provide a framework that can help researchers and instructors characterize and address the mathematical

TABLE I. A map of the different uses of the negative sign in elementary algebra, reproduced from Vlassis [11].

Unary (Structural signifier)	Symmetrical (Operational signifier)	Binary (Operational signifier)
Subtrahend	Taking opposite of	Completing
Relative number	or inverting the operation	Taking away
Isolated number		Difference between numbers
Formal concept of neg. number		Movement on number line

conceptualization of signed quantity in introductory physics.

In the next section, we describe the development of the NoNIP framework, including its basis in an analogous framework developed by mathematics education researchers. We present the NoNIP framework, along with examples of quantities and relationships that illustrate its use. We end Sec. II with a discussion of the validation of this framework in the context of introductory physics.

In Sec. III, we analyze recent studies in introductory and in upper-division physics using the NoNIP framework as a lens through which student cognition can be categorized and understood. Section IV describes our exploration of student understanding of “positivity” and the necessity of extending this work to signed quantities more generally. We discuss our conclusions and implications for instruction in Sec. V.

II. MODELING THE NATURES OF NEGATIVITY

In this section we discuss the need for and development of a framework to understand the uses of the negative sign in introductory-level physics. In Sec. II A, we include work by researchers that has directly influenced and guided the development of the NoNIP framework. In Sec. II B, we describe the process by which this framework was developed, expert validated, and, as a result of expert input, modified. We present the current version of the framework in Sec. II C.

A. Underpinnings

Our initial model for the natures of negativity in introductory physics was based on the natures of negativity in elementary algebra, described by Vlassis [11]. Vlassis summarized the work of mathematics education researchers, identifying three distinct algebraic natures of negativity. The first, referred to as the *unary* nature, describes situations in which a negative sign is used in close association with a single quantity, and includes the formal concept of a negative number (e.g., the number -5). Two additional natures signify mathematical operations: the *binary* nature describes various conceptualizations of the negative sign as it is used in subtraction (e.g., $5 - 3 = 2$), while the *symmetrical* nature describes use of the negative sign to invert (i.e., take the opposite of) a number or operation [e.g., the first negative sign in the expression $-(-5) = 5$]. Table I summarizes these algebraic natures.

We found that while most uses of the negative sign that arise in introductory physics could be categorized using the map synthesized by Vlassis, the nuances of the physics described by the mathematics were often lost. We also found ourselves tempted to represent the physical meaning of a negative sign attached to a single quantity (unary nature in Table I) using the categories intended for operations, as negative quantities in physics in some cases represent a *process* rather than just an amount. Moreover, it was difficult for physics experts to reach consensus when attempting to categorize some quantities and relationships. Both of these issues are exemplified by our struggle categorizing the meaning of the negative sign when used with the quantity mechanical work. Below, we describe three possible categorizations of mechanical work using Vlassis’s framework, pointing out both benefits and drawbacks for each categorization.

- *Unary*: For this categorization, the negative sign is associated with a single quantity (mechanical work). This is consistent with interpreting the meaning of the negative sign, but gives no insight, for example, into how the negative sign signifies something different with mechanical work than it does with electric charge.
- *Binary*: Here, the negative sign could be seen as a sign of “removal” or decrease, as a negative net work is associated with the decrease in mechanical energy of a system. However, this is interpreting the meaning of negative work, not the meaning of the negative sign when used with the quantity work; it associates the operation of removal with the process of doing negative work. There are other interpretations of negative work (e.g., considering work as a scalar product of force and displacement) for which this categorization is not sufficient.
- *Symmetrical*: A negative value for the work on a system can indicate that (positive) work is done *by* that system, rather than *on* the system, thereby “inverting” the meaning (i.e., $W_{\text{on B by A}} = -W_{\text{on A by B}}$). While physically correct, this is a limited interpretation.

The mathematical natures of the negative sign described by Vlassis provided insights into the meanings of negative signs in physics contexts. We concluded, however, that describing how mathematics represents *physics* quantities and relationships would require a different categorization than that used for pure algebra. Vlassis’s map is based in semiotics, the study of symbols and their meanings. While

we found this appropriate for describing meanings of negative signs in mathematics, we felt a blended processes framework would be more productive for interpreting meanings of the negative sign in the mathematics used in physics.

Conceptual blending theory (CBT) provides a framework for understanding the integration of mathematical and physical reasoning. This theory, formulated by Fauconnier and Turner, describes a cognitive process in which a unique mental space is formed by merging two (or more) separate mental spaces [14]. The blended space can be thought of as a product of the input spaces, rather than a separable sum. According to CBT, development of expert quantification in physics would occur not through a simple addition of new elements (physics concepts) to an existing cognitive structure (arithmetic), but rather through the creation of a new and independent cognitive space. This space, in which creative, quantitative analysis of physical phenomena can occur, involves a continuous interdependence of thinking about the mathematical and physical worlds.

The design of the NoNIP framework was further inspired by Sherin's work on symbolic forms, which posits that

... successful (physics) students learn to understand what equations say in a fundamental sense; they have a feel for expressions, and this understanding guides their work... from the point of view of improving instruction, it is absolutely critical to acknowledge that physics expertise involves this type of flexible and generative understanding of equations. We do students a disservice by treating conceptual understanding as separate from the use of mathematical notations [15].

We take the approach that a preliminary step in helping students develop a feel for expressions in a way that can productively guide their work is to understand better how expert conceptualization is organized, and thus characterize an expert's feel for expressions in a way that is useful both to education researchers and instructors.

We note that the physics contexts that are typically used as applications in a mathematics course are limited to one-dimensional kinematics. Nonetheless, mathematics education research into student reasoning with quantity helps to build a framework for thinking about mathematical objects in physics contexts. In one such example, Chiu reports on problem-solving interviews in which participants articulated how they conceptualize negative quantity when performing arithmetic with signed numbers. Half of Chiu's sample population was drawn from middle school students (for whom negative numbers was a fairly new concept), and the other half from post-secondary students in mathematics-intensive areas of study (undergraduate and graduate students from mathematics and engineering). The participants were asked to explain six arithmetic expressions (e.g., How do you make sense of $-5 + 8$?), most of which involved their application to contexts like the stock market or computing. Chiu identified three categories of

metaphorical reasoning: motion (movement along a number line), the manipulation of objects or opposing objects (removing or acting in opposition), and social transaction (associated with the experiences of giving and exchanging) [16]. The similarities in the reasoning patterns across Chiu's sample reveal that reasoning about negativity is strongly associated with first experiences in mathematics. While all participants relied on metaphors when they found the questions challenging, experts showed a more consistent, integrated understanding.

Whereas the metaphors revealed in Chiu's study are actual metaphors in pure mathematics, they are in fact *contexts* in physics. Conceptual mathematical understanding is essential for learning the physics of these contexts. Consider introductory-level mechanics, which is focused on actual motion in space, not metaphorical motion along an abstract number line. It is the interplay between physical quantities and their representation that motivates the creation of a negativity framework specific to physics, to help map out the trajectory to this expertise for all students—not just those who become majors.

B. First version of framework

To create a physics-specific framework of the natures of negativity, we began by generating a list of physics relationships involving an explicit negative sign (e.g., $\vec{F}_{\text{spring}} = -k\vec{r}$), and by considering base quantities (e.g., position, charge) and derived quantities (e.g., velocity, electric field) that can be implicitly negative or associated with a negative sign. Our intent was to develop a framework drawing on the practices and conventions of physics—which are situated in a blended space of mathematics and physics—to make meaning of negativity.

Learning scientists have used card-sorting tasks to investigate mental organization of disciplinary knowledge [17,18]. Experts are given cards showing various content with no pre-established groupings. They are then asked to sort the cards into groups that they feel make the most sense and describe each group. Two of us (S. W. B. and A. O.) employed a modified card-sorting task involving the quantities and relationships that make up the introductory physics course. On each card was a physical quantity that can be negative (e.g., one-dimensional velocity, temperature, potential difference) or a relationship between quantities involving an explicit negative sign. These quantities and relationships spanned the entire instructional sequence presented in the introductory textbook by Mazur [19]. The general sorting process was modified such that the researchers focused uniquely on the role of the negative sign for each quantity or relationship. Each researcher created groups, and the researchers discussed their respective groupings until both agreed to broad categorization based on physical similarities. Together, they gradually refined the groupings through further discussion and comparison with other quantities, resulting in

TABLE II. First version of the natures of negativity in introductory physics. The top row includes the four natures. The subnatures are organized into groups, shown bolded in the table. Some groups contain only one subnature.

(D) Direction	(O) Opposition	(Ch) Change	(Co) Compound
1. Location x	1. Opposite type Q (<i>charge</i>)	1. Removal (operator) $0 - (-5 \mu\text{C})$	1. Scalar rates of change $\frac{d\phi}{dt}$
2. Direction of motion $v_x, \Delta x$ p_x	2. Opposes $\vec{F}_{12} = -\vec{F}_{21}$ $\vec{F} = -\vec{\nabla}U$	2. Difference (operator) $U_f - U_i$ $\vec{p}_f - \vec{p}_i$	2. Base + change $\phi + \frac{d\phi}{dt} t$ $\vec{v} + \vec{a}t$
3. Other vec. quant. comp. E_x, B_x F_x, L_z a_x $\Delta p_x, \Delta v_x$	3. Scalar products $\mathcal{E} = -\frac{d\Phi_B}{dt}$ $\vec{F} = -k\vec{r}$ $W = \vec{F} \cdot \Delta\vec{x}$ $\Phi = \vec{B} \cdot \vec{A}$	3. System scalar quantities $\Delta K, \Delta U$ ΔS	3. Products $f(x)dx$ $E(r)dr$ $P(V)dV$
4. Above or below reference T (<i>temperature</i>) V (<i>electric potential</i>)		4. Scalar, vector change $\Delta U = U_f - U_i, \Delta V = V_f - V_i$ $\Delta \vec{p} = \vec{p}_f - \vec{p}_i$	4. Models $W_{\text{net,ext}} = \Delta U$ $\vec{F}_{\text{net}} = m\vec{a}$ $\Delta U = Q - W$

four natures of negativity specific to physics. Creation of a *change* nature of negativity exemplifies this process. The importance in physics of change and of conserved quantities led us to categorize many different and seemingly disparate uses of the negative sign in a nature related to change, because of an underlying connection to the calculation of and reasoning about change. This includes the negative sign as an operator (to calculate change or to signify the physical removal of a quantity from a system); the negative sign as an indication of a decrease in a quantity; and the compound use of the negative for signifying a change in a quantity as well as calculating it.

Along with change, the emergent natures were *direction* and *opposition*. A fourth nature, *compound* was added to account for cases that require interpretation of multiple negative signs in a single context. We note that the direction and opposition natures are supported by the categories isolated in Chiu's study [16]. Phenomena that arise due to the parallel or antiparallel orientations of two quantities are ubiquitous throughout physics (e.g., speeding up or slowing down, friction and air resistance, electromagnetic induction). Direction and opposition are central natures of signed quantities in physics, and hallmarks of physics reasoning.

To allow for further refinement, categories, or "subnatures" emerged within each of the main natures of negativity. Some categories also specify the mathematical function of the negative sign [as in the difference (operator) subnature within the change nature]. Table II shows the resulting map of the natures of negativity in introductory physics. We do not attempt to further explain the meanings of the categories shown in the table, as subsequent work led us to reorganize this early version of the NoNIP framework.

The first version of the NoNIP framework, Table II, was extensively validated. We assessed face validity by

surveying three, epistemologically varied introductory physics textbooks [19–21], using the NoNIP framework to categorize all instances of the use of negative signs. We found only one instance of use of the negative sign (the negative exponent) that did not fit into one of the categories of the NoNIP framework as we designed them. To provide additional validation of the NoNIP framework, we recruited experts in mathematics and physics with whom to conduct interviews.

We had two primary goals for the expert validation interviews: to ensure that the framework appropriately characterized the various meanings of the negative sign as used in introductory-physics contexts; and to ensure that there were no uses of the negative sign not adequately categorized by the NoNIP framework. We identified and recruited two experts each in mathematics and physics. Our physics experts are experienced introductory-level instructors; our mathematics experts taught introductory- and intermediate-level undergraduate mathematics (for example, single-variable and multivariable calculus and differential equations), with sufficient background in physics to understand physics-specific meanings (e.g., one of our mathematics experts has an undergraduate degree in physics).

Prior to the interviews, we provided our experts with a current version of the NoNIP framework, with accompanying text that described the purpose of the work, and the intended interpretations of the categories and subcategories. We asked our interview subjects to review the materials prior to their scheduled interview.

The interviews themselves were performed via video chat, lasted 30–60 min, and followed a formal, semistructured format. Two of us (S. W. B. and A. O.) performed the interviews together, with one researcher leading the interview and the other taking handwritten or typed notes. Each

TABLE III. Current version of the natures of negativity in introductory physics, a framework for the different uses of the negative sign in introductory physics. The top row includes the three natures. The subnatures are organized into groups, shown bolded in the table. Some groups contain only one subnature.

(Q) Quantity	(R) Relationship	(O) Operation
1. Scalar	1. Opposes	1. Removal (physical)
a. Type (charge only)	a. Scalar	$0 - (+5 \mu\text{C})$, $m_{\text{total}} - m_a$
b. Change or rate of change	$\mathcal{E} = -\frac{d\Phi_B}{dt}$	2. Removal (modeling)
$\Delta U, dV, \frac{d\phi}{dt}$	b. Vector	$I_{\text{net}} = I_{\text{disk}} - I_{\text{hole}}$
c. Comparison to reference	$\vec{F} = -\vec{\nabla}U$	3. Difference (change)
T, V, U, t	$\vec{F}_{\text{spring}} = -k\Delta\vec{x}$	$U_f - U_i$
d. Models or convention	2. Opposite	$\vec{p}_f - \vec{p}_i$
$W_{\text{net,ext}}$, Heat (Q), Current (i)	$+5 \mu\text{C} + (-5 \mu\text{C}) = 0$	4. Difference (other)
2. Vector	$\vec{F}_{12} = -\vec{F}_{21}$	1st law of thermodynamics ($Q - W$)
a. Direction from origin	3. Relative orientation	Path length difference ($D_2 - D_1$)
x, E_x, v_x	$\vec{F} \cdot \Delta\vec{x}, \vec{E} \cdot \vec{A}$	Separation ($\vec{r}_2 - \vec{r}_1$)
b. Direction of change	4. Negative exponents	Electric potential difference ($V_b - V_a$)
$\Delta p_x, \Delta v_x$	$e^{-t/\tau}, r^{-2}$	

interview began with a discussion of the materials sent to the expert, to ensure that the interview subject understood the purpose of the framework as well as the purpose of the interview. The experts were given the opportunity to ask for clarification of any aspect of the NoNIP framework. Interview questions were open ended and intended to inspire conversation about the framework.

Experts were asked to comment specifically on the appropriateness of the map in the context of introductory physics. The resulting conversations typically involved a discussion of the precise meanings of the nature and subnature categories, and how quantities or relationships would fit into those categories (or not). We found that the compound nature, in particular, was difficult for our experts to understand. Some experts deemed the nature unnecessary, stating a preference that we include a note that multiple negative signs with different meanings may be used in some cases. We also discussed, with each of our experts, whether there are uses of the negative sign in introductory physics that were not compatible with the NoNIP framework. Only one such instance was identified by any of our experts: negative exponents. The fact that this agreed with our own efforts contributes to the validity of the NoNIP framework. Mathematics experts were asked to comment on our interpretations of the negative sign from a more algebraic perspective. Each of these experts stressed the importance of recognizing the meaning of “zero” for a given quantity.

C. Steady state version of the NoNIP framework

While feedback from experts in interviews—and during less formal interactions during conference presentations—was positive, comments such as those described above led to a number of changes to the NoNIP framework. After each interview, the two researchers performing the interviews reviewed the notes, adding notes about possible changes to the framework. Some changes were made as

direct results of experts’ comments, while other changes occurred as our own understanding changed over the course of multiple interviews. These changes were initially small, but ultimately led to a reorganization. In particular, mathematics experts’ focus on the meaning of zero for different quantities led us to consider the negative sign in the unary sense (that is, attached to a single quantity) separately from other uses, and allowed us to distinguish more clearly between scalar and vector quantities. Comments by physics experts led us to consider when quantities are *opposite* one another [as -5 is opposite $+5$, such that $+5 + (-5) = 0$] and when quantities *oppose* one another (as in Faraday’s law, $\mathcal{E} = -d\Phi_B/dt$). Additionally, we recognized the importance of differentiating between the negative sign as an operator and its other uses. Finally, we made the decision to remove the compound nature entirely, following experts’ suggestions that we simply note that some contexts require the interpretation of multiple negative signs in one expression.

Reorganization of the NoNIP framework resulted in a framework with more algebraically-based natures. Each nature is further divided into multiple groups, with each group containing one or more subnature. For instances in which a group contains only one subnature, the subnature is fully described by the group—i.e., the group is the subnature. We believe that the major strength of both the initial and revised versions of the NoNIP framework is the physical interpretation present in the natures as well as the groups and subnatures.

The revised NoNIP framework is shown in Table III. We define three natures of negativity based on the functions of the negative sign:

- specifying its use with a single *quantity*,
- defining the *relationship* between two quantities, or
- as indicating the *operation* of subtraction.

The natures and subnatures give meaning specific to physics. We have removed the compound nature because it is not independent of these three: all examples from our original compound nature may be seen as combinations of subnatures of the three redefined natures. We discuss the importance and complexity of combining these natures of negativity more fully below. It is worth noting that some quantities (such as mechanical work and electric charge) appear in multiple subnatures. This speaks to the challenges that students face when trying to decode and make sense of negative signs (and implicit negative values) in introductory physics.

Like the original version of the NoNIP framework, this version was validated through individual, semistructured interviews, in this case with three physics experts. These experts are Ph.D. physicists who are experienced in teaching introductory courses and have above-average ratings in end-of-course student evaluations. All three experts have substantial knowledge and awareness of research-based, student-centered approaches to instruction, as evidenced by their teaching practice and their ongoing involvement in professional development workshops focused on pedagogy. The structure of the validation interviews was similar to that described above for the earlier version of the NoNIP framework.

All three experts indicated that the framework is complete, and captures the important ways that negative signs are used in undergraduate physics contexts. The experts further agreed that the framework includes thinking and understanding beyond what should be expected for students to master in a year-long introductory course. Two of the experts explained that some of the subnatures of negativity differ in ways that are quite subtle, and that would not likely be apparent to students even after a high-quality, year-long physics course. (One expert felt that a simplified version of the framework could be useful as a tool to present explicitly to introductory students.) The experts agreed that many of the natures are interrelated, with one expert explaining that the natures should not be regarded as an “orthogonal set” of ways that sign is used in introductory physics contexts. The experts differed somewhat in regards to vectors and scalars. One felt that a vector component and a scalar had little-to-no substantial difference, and stated that “coordinate axes” or “coordinate systems” should be explicitly mentioned in the framework, as these constitute a critical tool for understanding positive and negative natures of quantities in physics [22]. In contrast, the other two experts felt that scalars and vector components are clearly different types of quantities, with some differences in the associated meanings of a negative sign.

Based on these expert interviews, we regard the NoNIP framework to be validated as comprehensive and well-aligned with expert understanding of the meanings of the negative sign as used in introductory-level physics. Our experts did have varying opinions on some aspects of the framework; we deem these differences in opinion to be

relatively minor, and believe they do not justify modification of the steady-state NoNIP framework.

1. The quantity nature of negativity

The quantity (Q) nature of negativity is most similar to Vlassis’s unary nature (i.e., identifies a number as negative), with the negative sign attached to a single quantity. These quantities may be either positive or negative, and the sign of the quantity is often implicit, unless a specific negative numeric value is known. The quantity nature includes groups for 1. Scalar and 2. Vector quantities, which we treat differently due to the different meanings of the negative sign associated with each.

Four subnatures exist for scalar quantities.

- (Q1a) **Type:** sign acts as a label to specify the type of electric charge (or color charge in advanced courses).
- (Q1b) **Change or rate of change:** negativity indicates that a quantity is decreasing. This includes macroscopic changes (e.g., ΔU), scalar time rates of change (e.g., $d\phi/dt$), and differentials (e.g., dV) where a negative differential indicates an infinitesimal decrease in a quantity.
- (Q1c) **Comparison to reference:** a negative value is less than an arbitrary “0” reference point (e.g., temperature, electric potential, energy).
- (Q1d) **Models or convention:** sign carries important physical meaning that is an artifact (sometimes arbitrary) of a specific model. For example: heat Q is negative for a system when thermal energy is transferred out of that system; negative net work done on a system corresponds to a decrease in the mechanical energy of the system; and current i is negative when it is opposite to the sense arbitrarily decided to be positive.

We define two subnatures of negativity for vector quantities. In both subnatures, the sign of the vector represents a direction with respect to a defined coordinate system and origin.

- (Q2a) **Direction from origin:** sign indicates the direction of the component(s) of a single vector, e.g., position, velocity, or electric field.
- (Q2b) **Direction of change:** sign indicates the direction of the difference between two vectors, e.g., change in velocity or change in momentum.

2. The relationship nature of negativity

In the relationship (R) nature of negativity, the negative sign indicates how quantities *relate* to each other. This is most similar to Vlassis’s symmetrical nature (i.e., when a negative sign is used to take the opposite of or invert a number or operation); however, we do not consider the negative sign to be indicative of an *operation* on the quantities described in the relationship nature. While in mathematics a negative sign used to invert a number may be

seen as an operator that defines one quantity as opposite another, in physics the negative sign more generally indicates that two quantities are inversely related.

As with the quantity nature, we consider relationships between both vector and scalar quantities; however, the similarities and differences between these relationships vary between the uses of negative signs. For this reason we find it more productive not to distinguish between vector and scalar quantities as the first level of categorization for the relationship nature. A major difference from the quantity nature is that the negative sign is explicit in most subnatures in the relationship nature, with the exception being the scalar (dot) products in the relative orientation (R3) subnature.

- (R1) **Opposes:** equations with explicit negative signs to indicate the opposing relationships between quantities of different types. These relationships often involve derivatives of products or one side of the equation. A typical **Scalar** (R1a) example is Faraday's law, in which the EMF opposes the time rate of change of the magnetic flux. Similarly, we have **Vector** (R1b) relationships, such as Hooke's law, where the force exerted by a spring opposes the spring's displacement from equilibrium.
- (R2) **Opposite:** explicit negative signs indicate that one quantity is the opposite of another quantity of the same type. Examples include positive and negative charge, and the members of a Newton's third law force pair [23].
- (R3) **Relative orientation:** an implicitly negative scalar product—such as that used to calculate mechanical work ($\vec{F} \cdot \Delta \vec{x}$) or electric flux ($\vec{E} \cdot \vec{A}$)—indicates that the factor vectors have components that are oppositely oriented. We include scalar products in the relationship nature because, as with other expressions and quantities in this category, a negative value says something about how two quantities (in this case, the factor vectors) relate to each other.
- (R4) **Negative exponents:** negativity describes how one quantity relates to another, such as exponential decay describing how a quantity decreases in time, or a fixed negative exponent describing how a quantity varies in space (e.g., r^{-2} or $1/r^2$).

Recognizing the equivalence between a quantity with a negative exponent and the inverse of that quantity with a positive exponent is vital and (in our experience) nontrivial. This addresses a major shortcoming that existed in the original version of the NoNIP framework, which did not allow for straightforward categorization of negative exponents. More advanced applications, such as determining solutions to boundary-value problems, emphasize the relationship nature of negative exponents: explicit negative signs are often used to indicate qualitatively different solutions. For example, the solution to the differential equation,

$$\frac{d^2f}{dr^2} = n^2 f(r) \quad (1)$$

is often written as

$$f(r) = \sum_{n=0}^{\infty} A_n e^{nr} + B_n e^{-nr}, \quad (2)$$

with the integer parameter n assumed to be positive. It is not strictly necessary to include both terms in the summation (the summation could be taken from $-\infty$ to $+\infty$ instead with only the leading term), but the explicit negative sign allows a convenient separation between groups of solutions that have very different behavior at large (and small) values of r .

3. The operation nature of negativity

The operation (O) nature of negativity includes instances when the negative sign is used to perform the mathematical function of subtraction, which is similar to Vlassis's binary nature. By necessity, the minus sign is always explicit in the operation nature [24]. We define four subnatures of the operation nature, with two overarching themes: Removal of a part from a whole, and taking the difference between two quantities.

- (O1) **Removal (physical):** the minus sign signifies a physical removal of some quantity from a system (such as electric charge or mass).
- (O2) **Removal (modeling):** the nonphysical removal of one quantity from another for the purposes of modeling a more complex situation. This type of subtraction is typified by the calculation of the moment of inertia of a thick ring modeled as a solid disk that is missing another solid disk (a hole).
- (O3) **Difference (change):** calculating the temporal change in a quantity, such as the change in energy $U_f - U_i$ or change in momentum $\vec{p}_f - \vec{p}_i$.
- (O4) **Difference (other):** an operation used to calculate differences that are not temporal in nature: the subtraction of work done by a system from the heat supplied to the system in the first law of thermodynamics, or the separation between two point charges in Coulomb's law.

We find it necessary to include changes (temporal differences) in both the quantity and operation natures because a negative sign associated with the quantity ΔU indicates that the final value of the energy is less than the initial value, but the minus sign in category O3 represents the operation of subtracting one quantity from another (without providing any information about their relative values).

Carefully examining the similarities and differences between subnatures O3 and O4 highlights the complexity of common notation. A typical example of difference (other) is the calculation of displacement for Hooke's

law, with $\Delta\vec{x} = \vec{x} - \vec{x}_{\text{eq}}$. In this case the “ Δ ” notation does not represent the change in position $\Delta\vec{x} = \vec{x}_f - \vec{x}_i$ (which would show up in both O3 and Q2b), but the displacement from equilibrium. Expert physicists are likely to understand the implicit distinction between the meanings of these notations, but the subtleties may be difficult for students to recognize.

4. Compounding multiple natures of the negative sign

Expert feedback led us to recognize that creating a compound nature in the original NoNIP framework somewhat hid the cognitive difficulty students encounter unpacking the multiple natures of negativity from a single expression or equation. Because of the many possible combinations of signed quantities appearing in an expression, each compound context poses a unique challenge. It is in these compound contexts that even very strong students struggle most. A significant difference between the revised version of the NoNIP framework (Table III) and the original version (Table II) is the new version’s lack of a compound nature. We believe that quantity, relationship, and operation represent the individual natures of negativity in introductory physics thoroughly. The task of combining them in compound cases is an example of the sophisticated reasoning with familiar mathematics that is characteristic of well-developed quantitative literacy.

Many models in physics involve more than one unique nature of negative quantities; we consider the compound expressions ubiquitous in physics to be incomprehensible without flexibility between the three main natures of negativity. Figure 1 maps two examples to the current NoNIP framework, which reveal various uses of implicit and explicit negatives from all three natures of negativity that must be understood in order to make physical sense of central ideas in introductory physics. One must note that some quantities (such as electric charge) appear under multiple natures, as the sign carries multiple meanings. Moreover, in some cases, such as Coulomb’s law, multiple negative signs may “cancel out.” Such “hidden” negative signs make keeping track of the signs of individual components more challenging. When Coulomb’s law is used with two charges with *opposite* signs, students must also make sense of what the resulting negative sign implies. For physics experts, a negative sign implies that the force between the two charges is attractive, but this meaning is also described by the mathematics: that $\vec{F}_{\text{by } 1 \text{ on } 2}$ is oppositely directed to $\vec{r} = \vec{r}_2 - \vec{r}_1$. We contend that this physical interpretation of a quantified relationship is a hallmark of expertlike reasoning.

The identification of multiple meanings of negative signs in a single context is challenging and the ability to do so is associated with expertlike thinking. It should not be considered to be a skill at the same level as identifying the meaning in less complex cases; rather, it should be viewed as a culmination of reasoning with and about sign.

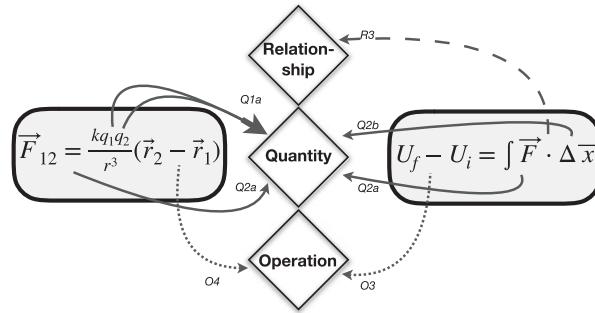


FIG. 1. Compound use of negative signs typical of introductory physics: Coulomb’s law, the force on object 2 through its interaction with object 1 (left), and the work-energy theorem (right).

We argue that flexibility between the various natures that appear in the NoNIP framework may better prepare students for the challenge of combining them in a single equation. Indeed, much of the research on more advanced student reasoning about sign has been in the context of compounded use of the negative sign. In the following section, we describe such research, using the NoNIP as a framework by which student difficulties can be understood. In doing so, we provide examples of how to use the 3-nature NoNIP framework to categorize compound quantities and relationships.

III. APPLYING THE FRAMEWORK

In this section, we use the current, steady-state version of the NoNIP framework (Table III) as an analytical lens through which to view several recently published studies of student difficulties in physics and calculus. As a proof of concept that the NoNIP framework characterizes novice or expert gaps at the introductory level, which are not unique to universities in the United States, we demonstrate the utility of the framework in two studies involving introductory physics contexts from three different continents. We then use the NoNIP framework to characterize difficulties in multiple contexts in advanced physics courses in the United States, demonstrating the persistence of these difficulties—even for majors.

Brahmia and Boudreaux investigated calculus-based introductory physics students’ interpretation of the meaning of the negative sign in mechanics and electricity and magnetism contexts in a large research university in the northeast United States [25]. The researchers analyzed free-response questions administered at the end of the semester (either mechanics or electricity and magnetism) in which students were asked to interpret the meaning of the negative sign in the one-dimensional vector quantities $a_x = -8 \text{ m/s}^2$ or $E_x = -10 \text{ N/C}$, depending on which course they were taking. The expected response is that the negative sign indicates that the vector quantity points in the negative direction. The researchers found that the majority of students misinterpreted the meaning by attributing additional physics meaning to the sign. In the mechanics

course, students typically associated the sign with features of the direction of motion or speed, conflating the change in velocity direction with the direction of the velocity itself, and in the electricity and magnetism course they associated it with the nature of the charges involved. Imputing physics meaning to the mathematical object involved can be characterized in Table III as conflating Q2b with Q2a, and Q2a with Q1a, respectively.

In their study of introductory physics students at universities in Sweden and South Africa, Eriksson, Linder, and Eriksson found that many students viewed the sign of velocity in one-dimensional motion as a description of how the motion *changed*, i.e., either speeding up or slowing down [9]. Confusing the direction of a one-dimensional vector quantity with an increase or decrease of its magnitude during a time interval is represented as conflating Q2a with O3.

Bajracharya, Wemyss, and Thompson investigated upper-division student understanding of integration in the context of definite integrals commonly found in introductory physics, but with all physics context stripped from the representation. Specifically, the variables typically used in physics contexts were replaced with x and $f(x)$ [7]. Their results suggest difficulties with the criteria that determine the sign of a definite integral. Students struggle with the concept of a negative area under the curve, and in particular negative directions of single-variable integration. In research related to student understanding of integration and negativity, Sealey and Thompson interviewed undergraduate and graduate students to uncover how they made sense of a negative definite integral. Undergraduate (beyond introductory) and graduate mathematics students had difficulty making meaning of a negative differential in the context of integration [26]. The struggles these researchers described can be seen through the lens of the NoNIP framework as struggle with the product of the integrand, $f(x)$ (generic Q in NoNIP), and the differential, dx (Q1b in NoNIP), each of which can independently be negative. Making meaning of the negativity of the integrand (generic Q in NoNIP) was less of a struggle for the students in these studies than was the notion of a negative differential (Q1b in NoNIP), which has application throughout physics. The researchers report that “none of the students thought about dx as a signed quantity on their own accord, but with prompting from the interviewers, some were able to do so.” Encountering the differential as a small change in quantity, and being provided opportunity to think about it in this way, provides a context that has been shown to help [27]. The NoNIP framework can support this kind of explicit reasoning about sign in this important context.

A study conducted by Hayes and Wittmann situated in the context of a sophomore-level mechanics course investigates the negative signs and quantities associated with the equation of motion of an object thrown downward, with non-negligible air resistance [6]. The student described in their study (representing the answers of roughly half the

class) struggles with treating one-dimensional acceleration as a signed net effect, and feels there should be an additional negative sign included to indicate that the acceleration is “negative,” or opposing the motion. The student writes $-ma = mg - cv$. The negative sign on the right-hand side of the equation collapses the operation of vector addition used to determine a net force with the negative sign in front of the resistive force, which indicates that that force opposes the velocity. The student imputes a rule about the signs to encode physical meaning, misattributing the sign on the right-hand side to the direction of the velocity, and includes an extra sign for the acceleration. Viewed through the lens of the NoNIP framework, the student struggles with Q2a in the contexts of one-dimensional acceleration and velocity. The negative sign that modifies the cv term is used as R1b, to indicate that the resistive force is in the opposite direction to the velocity. In the process of combining these terms, the student struggles to make sense of the equation of motion, where the direction of the acceleration is the *result* of the vector addition. The cognitive load of negativity associated with the individual terms contribute to a higher-level struggle of making physical sense.

In their study of negativity in junior-level Electricity and Magnetism, Huynh and Sayre describe the in-the-moment thinking of a student solving for the electric field due to two equal and opposite charges along the axis that passes through them [8]. The authors focus on student reasoning about the sign of the electric field vector component along the axis of symmetry in three regions of space—to the left of one charge, between the two charges, and to the right of the other charge. They report that four students solve this problem in an oral exam, and none get the directions correct on their first attempt.

The solution involves an algebraic superposition of the field due to each charge individually. The authors detail one representative student’s development of an increasingly blended approach that is situated in a mental space informed by both mathematical and physical concepts. The student starts reasoning about the direction of the field using Coulomb’s law by (unintentionally) combining multiple natures of negativity into one. In Coulomb’s law, the signs combine multiplicatively from two sources: the direction of the vector, $\vec{r}_2 - \vec{r}_1$, and the sign of the source charge (see Fig. 1). The student first uses the canceling feature that multiplying two negative numbers always results in a positive number, without explicitly considering the source of each negative sign, and then reflects based on physics considerations why that approach does not make sense.

The student considers \hat{x} to be a proxy for \hat{r} , without considering that \hat{r} is connected to the physics (related to the difference of the two position vectors) while \hat{x} is a feature of the coordinate system. By making this substitution, the student glosses over an important source of negativity in the final solution and it keeps him from being able to calculate

the direction of the electric field that he predicts using physics principles. His conceptual understanding of the physics is strong, but he cannot make his calculation match because he is not considering that the source of the negative signs have deep physical meaning beyond the charges involved. Seen through the lens of the NoNIP framework we can see evidence of the student first conflating Q2b and Q2a, not recognizing that they are not the same thing. The authors summarize the student's confusion regarding the relative signs of the contributions to the electric field of each charge individually:

This result clearly conflicts with their relative direction because he has double associated their opposite direction with inappropriate application of destructiveness. [The student] tries hard to determine where another negative sign could come from, such as the denominator, to cancel one negative sign for the whole term. Finally, he decides to absorb the destructive meaning of the sign into the opposite-direction meaning of the electric field vector and changes the second negative sign of the whole term back into a plus sign, which supports the fact that they are in opposite directions. However, [the student] has not considered the sign commensurate with the relative direction of \hat{E} and \hat{x} , leading to his solution having the opposite sign to the correct answer [8].

Collapsing the signs using arithmetic rules is a common first approach of the students in this study, which focuses on the multiplicative rules of signed numbers rather than the physics of the meaning of the signs. Next, the student rarefies his approach as he considers more carefully the natures of negativity in the context of the problem. The student reflects "... I should have figured it out ... which direction it is. This is exactly what is changing signs, not necessarily the sign of the charge." The student uses conceptual physics reasoning consistent with R1b, then he struggles with O4 in generating a mathematically consistent argument. The authors report that when the student moves on to the other two regions, which is essentially repeating the same reasoning sequence, the student encounters the same struggles but he is faster at obtaining a solution that matches his physical understanding of the system. The fact that he does not automatically and quickly solve the remaining two regions is evidence that this kind of reasoning is difficult.

The authors conclude, and we agree, that the most sophisticated challenge occurs when the natures of negativity are combined—the compound nature presents its own challenge in addition to the challenge associated with each nature individually. This case study reveals the cognitive difficulty when three natures of the negative sign must be made sense of in the context of a single equation, and illustrates the challenges associated with reasoning about the natures of negativity, even for strong majors.

We believe it also reveals a hierarchy that lends plausibility to the NoNIP framework being representative of emergent expertlike reasoning.

This section, in which previously published work is interpreted through the NoNIP framework, demonstrates the benefits of the framework: it allows comparisons across studies that involve different physics contexts and even different mathematical approaches. Both the Huynh and Sayre and the Hayes and Wittmann studies involve upper-division topics and situations involving negative signs [6,8]; otherwise there are few similarities. The Hayes and Wittmann paper is situated in classical mechanics and difficulties with "inner and outer" negative signs, while the Huynh and Sayre paper investigates student ability setting up integrals in electricity and magnetism. By using the NoNIP framework, we can see that the two papers discuss similar aspects of the use of the negative sign. These similarities are much more specific than just using a negative sign in upper-division contexts. Further, the study by Bajracharya, Wemyss, and Thompson [7] is also about using negative sign in an upper-division context, but applying the NoNIP framework reveals a use of a different nature of negativity than the other two studies. We believe this illustrates the power of the NoNIP framework for researchers: identifying similarities across (and differences between) studies that are more than just superficial.

IV. EXTENSION TO SIGNED QUANTITIES

Although we focus in this paper on our categorization of the negative sign in introductory physics, we recognize that students must make sense of the meaning of positive quantities and relationships as well. Our focus on the negative sign in this work is due to the assumption of positivity when a quantity has no explicit sign. While the practice of assuming an unsigned number has an implicit positive sign in front of it perhaps poses few problems understanding pure numbers, quantities in physics that are not signed may be either unsigned scalars (e.g., speed, mass), or quantities in which the positive sign holds meaning (e.g., component of velocity, change in energy).

To investigate student understanding of signed quantities more generally, we administered three questions about positive or negative quantities in a multiple-choice format, shown in Fig. 2, to students enrolled in a calculus-based introductory physics course at a large, diverse, public R1 university. Each student received either all three negative or all three positive versions. Figure 3 shows the results from the negative and positive versions of the mechanics items ($N_{\text{pos}} = 242$, $N_{\text{neg}} = 309$).

To determine the effect size, we use the odds ratio. It is a useful effect-size measure that describes the likelihood of an outcome occurring in the treatment group compared with the likelihood of the outcome occurring in the control group by forming a ratio of the two. An odds ratio of 1 would indicate that the odds are exactly the same. If we

ME1: An object moves along the x-axis, and the acceleration is measured to be $a_x = +8 \text{ m/s}^2$. Consider the following statements about the “+” sign in “ $a_x = +8 \text{ m/s}^2$ ”. Pick the statement that best describes the information this positive sign conveys about the situation.

- The object moves in the positive direction
- The object is speeding up.
- The object accelerates in the $+x$ -direction**
- Both a and b
- Both b and c

ME2: A hand exerts a horizontal force on a block as the block moves on a frictionless, horizontal surface. For a particular interval of the motion, the work W done by the hand is $W = -2.7 \text{ J}$. Consider the following statements about the “-” sign in the mathematical statement “ $W = -2.7 \text{ J}$.” The negative sign means:

- the work done by the hand is in the negative direction
- the force exerted by the hand is in the negative direction
- the work done by the hand decreases the mechanical energy associated with the block

Which statements are true?

- I only
- II only
- III only**
- I and II only
- II and III only

ME3: A cart is moving along the x-axis. At a specific instant, the cart is at position $x = -7 \text{ m}$. Consider the following statements about the “-” sign in “ $x = -7 \text{ m}$.” Pick the statement that best describes the information this negative sign conveys about the situation.

- The cart moves in the negative direction
- The cart is to the negative direction from the origin**
- The cart is slowing down
- Both a and b
- Both a and c

FIG. 2. Examples of a multiple-choice questions probing student understanding of signed quantities. ME1 is an example of a positive-quantity question, while ME2 and ME3 are negative-quantity questions. Correct answers are bolded.

consider the hypothesis that positive quantities pose fewer challenges for students than negative ones, we can consider the positive questions as the control and the negative questions as the treatment.

For items ME2 and ME3, effect sizes determined from odds ratios are 1.2 and 1.1, respectively, which imply that students struggle with both positive and negative versions with roughly equal likelihood on these questions. The effect size for item ME1 is 0.74, which is a statistically small effect size indicating that students find the negative version of this question slightly more difficult than the positive version [28]. Our experience as instructors has led

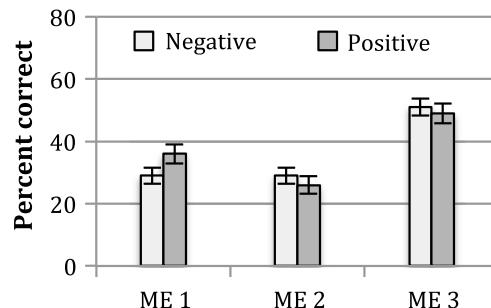


FIG. 3. Percentage of students who answered correctly for positive and negative versions of mechanics items, $N_{\text{pos}} = 242$, $N_{\text{neg}} = 309$; the error bars represent the binary standard error.

us to recognize that students tend to inappropriately associate negative (positive) acceleration with decreasing (increasing) speed regardless of the coordinate system. We suspect that the small difference in student performance on the two versions of item ME1 is related to this difficulty, and thus we do not interpret the finding as evidence that students experience inherent difficulty with negatively signed quantities (i.e., relative to positive quantities).

These results suggest that students have difficulty interpreting the meaning of the sign of a quantity, regardless of the sign; students may not recognize that the sign specifies the direction of a vector component relative to a coordinate system, or that the sign of a scalar quantity such as work indicates how the energy of a system changes.

Informal conversations with students also indicate that students sometimes fail to see the significance of the relationship between two quantities that are positively correlated (e.g., that Newton’s second law tells us that the acceleration of a system is always in the same direction as the net force exerted on that system). Such an understanding is crucial not only for nominally causal relationships, but also for understanding feedback loops and accumulated change. While interpreting negative signs is a pressing issue in physics learning because negative signs are explicitly used, these difficulties fall under the umbrella of student difficulties with the interpretation of signed quantities in general—which has not been studied in depth.

V. CONCLUSION

Negative signs in physics have nuanced and varied interpretations that can pose a challenge, even to majors. In this paper, we present the NoNIP framework for categorizing expert uses of the negative sign which has undergone expert validation and revision; it is presented in its steady state. We anticipate that the NoNIP framework can be useful across the physics education research community. For researchers, the NoNIP framework can serve as a map of the natures of negativity, and a starting point for thinking about sign, as part of the broader context of mathematical reasoning development in physics context.

We have demonstrated how viewing published research through the NoNIP framework can help bring out patterns between the findings that were not clear before.

We present the results from one study that suggest that students struggle to make sense of positive signs as well as negative signs. Student difficulties with positivity are not as noticeable because in practice experts assume the absence of sign means the quantity is positive, so, unlike a negative quantity, there is no symbol there to decode. We intend for the research presented in this paper to extend to signed quantities more generally, and increase awareness that students will benefit from making sense of the meaning of positive quantities as well as negative. The sign of a quantity, along with the magnitude of the quantity and its units, are part of what defines a quantity and how we understand it in most physics contexts [29].

We believe the framework can be useful in guiding the design of physics instruction and physics teacher preparation. Currently, we are pilot testing an online HW-based intervention for an introductory physics course in collaboration with, and informed by the work of Mikula and Heckler [30]. The intervention involves multiple-choice training items that require students to distinguish different natures of negativity in physics contexts (e.g., to distinguish the meaning of a negative sign associated with a velocity component from the meaning of a negative sign associated with a decrease in temperature). Item choices are derived from different natures of negativity, only one of which is relevant to the specific situation presented in the item.

As a tool for instructional development, the NoNIP framework can benefit both curriculum developers—who can use the NoNIP framework to help guide their efforts to situate signed quantity reasoning in the broader context of the materials that they develop—and instructors. We close by presenting some recommendations that can inform instructors about the organization of their expertise, which can thus influence how they talk about and present material to the novices in their classroom. Developing this awareness can help students become more cognizant of the natures of both positivity and negativity in physics. Acknowledging the nuances, rather than assuming the mathematics to be trivial, can create access for students that otherwise might not exist. We offer three suggestions as a start, fully anticipating that expert instructors will devise their own ways also:

- (1) Quantities that are inherently signed quantities should be prefaced with a negative sign when the quantity is negative, and a positive sign when the quantity is positive, e.g., $x_0 = +40$ m, at least until students understand the meaning in context of the absence of a “+” sign. Priming students to expect that real-world quantities often have signs associated with them that carry meaning, and that “no sign” is a different kind of quantity than a positively signed quantity, can help establish a physics habit of mind

that the sign carries scientific meaning, and eventually that vector quantities have different mathematical properties than scalar ones.

- (2) Orientation (along a particular axis) and sense (positive or negative) are not always explicit in coordinate systems. In problems associated with motion, aligning the positive coordinate axis with the direction of motion eliminates the need for signed quantities when discussing velocity. This choice, however, could be a missed opportunity to distinguish between orientation and sense. The opposite coordinate choice can prime students to consider the signed nature of position, velocity, and subsequent vector quantities they encounter.
- (3) Sign and operation are often conflated using an equals sign [e.g., $5 + (-3) = 5 - 3$], and unsigned numbers are assumed positive. Adding a negative quantity and subtracting a positive one often have different meanings in physics contexts (e.g., adding electrons). Although these operations yield the same arithmetic results, conflating them may lead students to struggle with the distinctions between sign and operation. We suggest using the term “minus” for the operation of subtraction, and the term “negative sign” to describe the symbol.

In addition to enriching subsequent physics learning, a focus on natures of sign in physics contexts can also enrich the corequisite mathematics learning. Sealey and Thompson report on a context in which physics helps mathematics students make sense of negativity in calculus. The researchers observed that invoking a physics example of a stretched spring helped catalyze sensemaking—the physics helped students to make sense of an abstract binary nature of the negative sign [26]. We suggest that there is symbiotic cognition possible in which both mathematics and physics learning can be enriched by conceptualization of the other, and that reasoning about sign provides a rich context. We present the NoNIP framework as a representation of the natures of negativity providing a step in that direction.

ACKNOWLEDGMENTS

We thank Peter Shaffer for his support with data collection that informed this work, and the entire Physics Education Group at the University of Washington for vibrant discussion and feedback. Brian Stephanik’s feedback and physics content expertise was also instrumental in the construction of the NoNIP framework. We thank Roy Montalvo at Rutgers University for sharing his creative software innovation that helped to make the data collection run smoothly. We also thank Laurie Smith for fruitful conversations regarding negative exponents. We gratefully acknowledge the support of the National Science Foundation for making this work possible through the Grants No. DUE-1832836, No. DUE-1832880, and No. DUE-1833050.

[1] J. P. Bishop, L. L. Lamb, R. A. Philipp, I. Whitacre, B. P. Schappelle, and M. L. Lewis, Obstacles and affordances for integer reasoning: An analysis of children’s thinking and the history of mathematics, *J. Res. Math. Educ.* **45**, 19 (2014).

[2] A. Gallardo and T. Rojano, School algebra. Syntactic difficulties in the operativity, in *Proceedings of the XVI International Group for the Psychology of Mathematics Education, North American Chapter, Vancouver, Canada*, Vol. 1, 265–272 (1994).

[3] P. W. Thompson and T. Dreyfus, Integers as transformations, *J. Res. Math. Educ.* **19**, 115 (1988).

[4] S. Brahmia and A. Boudreux, Exploring student understanding of negative quantity in introductory physics contexts, in *Proceedings of the 19th Annual Conference of RUME, Pittsburgh, PA* (2016), p. 79.

[5] S. Brahmia and A. Boudreux, Signed quantities: Mathematics based majors struggle to make meaning, in *Proceedings of the 20th Annual Conference on RUME* (San Diego, CA, 2017), p. 1158.

[6] K. Hayes and M. C. Wittmann, The role of sign in students’ modeling of scalar equations, *Phys. Teach.* **48**, 246 (2010).

[7] R. R. Bajracharya, T. M. Wemyss, and J. R. Thompson, Student interpretation of the signs of definite integrals using graphical representations, *AIP Conf. Proc.* **1413**, 111 (2012).

[8] T. Huynh and E. C. Sayre, Blending mathematical and physical negative-ness, [arXiv:1803.01447](https://arxiv.org/abs/1803.01447).

[9] M. Eriksson, C. Linder, and U. Eriksson, Towards understanding learning challenges involving sign convention in introductory level kinematics, in *Proceedings of the 2018 Physics Education Research Conference, Washington, DC* (AIP, New York, 2018).

[10] S. Ceuppens, L. Bollen, J. Deprez, W. Dehaene, and M. De Cock, 9th grade students’ understanding and strategies when solving $x(t)$ problems in 1D kinematics and $y(x)$ problems in mathematics, *Phys. Rev. Phys. Educ. Res.* **15**, 010101 (2019).

[11] J. Vlassis, Making sense of the minus sign or becoming flexible in ‘negativity’, *Learn. Instr.* **14**, 469 (2004).

[12] S. White Brahmia, Negative quantities in mechanics: A fine-grained math and physics conceptual blend?, in *Proceedings of the 2017 Physics Education Research Conference* (AIP, New York, 2017) pp. 64–67.

[13] A. Olsho, S. Brahmia, T. Smith, and A. Boudreux, When negative is not “less than zero”: Electric charge as a signed quantity (to be published).

[14] G. Fauconnier and M. Turner, *The Way We Think: Conceptual Blending and the Mind’s Hidden Complexities* (Basic Books, New York, 2008).

[15] B. L. Sherin, How students understand physics equations, *Cognit. Instr.* **19**, 479 (2001).

[16] M. M. Chiu, Using metaphors to understand and solve arithmetic problems: Novices and experts working with negative numbers, *Math. Think. Learn.* **3**, 93 (2001).

[17] M. T. H. Chi, P. J. Feltovich, and R. Glaser, Categorization and representation of physics problems by experts and novices*, *Cogn. Sci.* **5**, 121 (1981).

[18] A. Schoenfeld and D. J. Herrmann, Problem perception and knowledge structure in expert and novice mathematical problem solvers, *J. Exper. Psychol.: Learning Memory Cognit.* **8**, 484 (1982).

[19] E. Mazur, C. H. Crouch, D. Pedigo, P. A. Dourmashkin, and R. J. Bieniek, *Principles & Practice of Physics* (Pearson, Upper Saddle River, NJ, 2015).

[20] P. A. Tipler and G. Mosca, *Physics for Scientists and Engineers* (Macmillan, New York, 2007).

[21] E. Etkina, M. Gentile, and A. Van Heuvelen, *College Physics* (Pearson Higher Ed., Upper Saddle River, NJ, 2013).

[22] This expert did not feel that the different behaviors of scalars and vector components under coordinate transformation was particularly important for the natures of negativity. (“Behaviors” here refer to a scalar remaining invariant and 1D vector component changing sign under a 180° rotation of the coordinate axis).

[23] It was only after much deliberation that Newton’s third law force pairs were categorized as “opposite” and *not* as “opposes.” Although it would be reductive to characterize all of the relationships in the opposes category as *causal*, many of those relationships are at least *directional*. Members of a third law force pair are not causal or directional in that fashion, despite the convention of calling them “action-reaction” forces. Further, the forces in a third law force pair act, by definition, on separate objects.

[24] In terms of vocabulary, we argue that the word “minus” should only be used in the context of *operations*, and that any *quantities* and *relationships* should be described using the word “negative.”

[25] S. Brahmia, A. Boudreux, and S. E. Kanim, Obstacles to mathematization in introductory physics, [arXiv:1601.01235](https://arxiv.org/abs/1601.01235).

[26] V. Sealey and J. Thompson, Students’ interpretation and justification of “backward” definite integrals, in *Proceedings of the 19th Annual Conference on RUME, Pittsburgh, PA* (2016).

[27] D. Roundy, T. Dray, C. A. Manogue, J. F. Wagner, and E. Weber, An extended theoretical framework for the concept of the derivative, in *Proceedings of the 18th Annual Conference on Research in Undergraduate Mathematics Education* (San Diego, CA, 2015), pp. 838–843.

[28] J. M. Maher, J. C. Markey, and D. Ebert-May, The other half of the story: effect size analysis in quantitative research, *CBE-Life Sci. Educ.* **12**, 345 (2013).

[29] S. White Brahmia, Quantification and its importance to modeling in introductory physics, *Eur. J. Phys.* **40**, 044001 (2019).

[30] B. D. Mikula and A. F. Heckler, Framework and implementation for improving physics essential skills via computer-based practice: Vector math, *Phys. Rev. Phys. Educ. Res.* **13**, 010122 (2017).