PHYSICAL REVIEW PHYSICS EDUCATION RESEARCH **19,** 010108 (2023)

Investigating the efficacy of attending to reflexive cognitive processes in the context of Newton’s second law

J. Caleb Speirs[*] and Robyn Leuteritz
School of Mathematical and Physical Sciences, University of New England, Biddeford, Maine 04005, USA

Thanh K. Lê
Department of Physics and Astronomy, Western Washington University,
Bellingham, Washington 98225, USA

Rose Deng and Shawn W. Ell
Department of Psychology, University of Maine, Orono, Maine 04469, USA

(Received 30 August 2022; accepted 3 January 2023; published 14 February 2023)

Even after research-based instruction, students who demonstrate the ability to assemble relevant conceptual knowledge on one physics question may have difficulty assembling that same knowledge on a closely related problem. Recent research has suggested that reflexive, bottom-up reasoning processes seemingly unrelated to the physics concepts themselves may be responsible for these difficulties. Research has also suggested that attending to these reflexive processes during instruction may improve performance to a greater degree than attending solely to top-down, reflective thinking. Leveraging these findings to meaningfully improve instruction is important. We have, therefore, investigated the impact of training focused on Newton’s second law targeted at reflexive reasoning processes and compared results to a more standard reflective approach to the same topic. We find that an approach targeted toward reflexive reasoning processes improves performance on a difficult physics question to the same or greater degree as a typical reflective approach. Furthermore, we find that many students whose performance on a difficult physics question increased after the reflexive training also explained correct conceptual reasoning on that question, suggesting that conceptual understanding was bolstered by the bottom-up, reflexive training.

DOI: 10.1103/PhysRevPhysEducRes.19.010108

I. INTRODUCTION

Physics education research seeks to improve the teaching and learning of physics. Oftentimes, the efficacy of physics instruction is measured by examining students’ ability to reason with physics concepts. Studies of human reasoning suggest two types of processes by which humans reason and make judgments: an implicit, reflexive process (process 1) and an explicit, effort-full, and often reflective process (process 2) [1,2]. Most research-based curriculums seemingly target reflective type 2 processes by asking students to think explicitly about physics concepts and test hypotheticals in various ways [3–8]. However, recent research has suggested that attending to reflexive type 1 processes in curriculum development can be beneficial [3,9–12]. This paper examines the impact of two activities designed to teach students to use Newton’s second law in the context of static friction: one that targets reflexive processing and one that targets reflective processing. In doing so, we attempt to explore the advantages of leveraging the reflexive processes to improve the learning of physics. Additionally, we aim to contribute to a broader understanding of the interplay between the two types of processes.

II. BACKGROUND

This paper uses dual-process theories of reasoning and decision making—a collection of theories drawn from cognitive and social psychology—as a theoretical framework [1]. In dual-process theories of reasoning (DPTOR), reasoning occurs through the interplay of two types of processes: the first type of process (process 1) typically occurs reflexively and is usually (but not always) associated with speed; the second type of process (process 2) is an effortful, explicit, and often reflective or algorithmic process [1,2]. As an illustration of the two types of processes, consider a child who is crying. Process 1 takes in the visual and auditory stimuli and automatically understands that the child is upset. Process 2 may engage to try to understand why the child is upset. It may cause the person

---

[*] Corresponding author.
j.caleb.speirs@unf.edu

to look more closely at the context and pursue multiple hypotheticals—did the child fall over? Did a beloved toy get taken away unexpectedly?

Dual-process theories of reasoning provide a framework for understanding many of the reasoning errors that students make in response to physics questions. For instance, Heckler and Scaife [13] investigated how a fast process 1 can produce answer patterns that had previously been described as *slope and height confusion* [12,14,15]. In this study, the working hypothesis was that slope and height confusion arises from a competition for attention between the two dimensions of slope and height coupled with the fact that the height (or value) of a point on a graph has a faster processing time relative to the time it takes to process the slope of a point on a graph. Because the height was processed faster, it won the students' attention. This can therefore be thought of as a fluency-based bias. They demonstrated reduced slope and height confusion errors in student responses by placing a mandatory waiting period (slightly longer than the average slope-processing time) after seeing the graph but before having the ability to answer. Thus, a process 1 mechanism (and not a process 2 mechanism alone) was shown to be responsible for student reasoning errors. Additionally, process 1 mechanisms have been associated with errors in reasoning about relevant factors in physical phenomena [16], the interpretation of data [17], the determination of the location of the center of mass [3,18], and the summation of vector quantities [19].

Dual-process theories of reasoning also provide insight into the structure of conceptual knowledge. In physics education research, conceptual understanding is largely assessed through the analysis of how students respond to various physics tasks [20–23]. In other words, student response patterns form the evidentiary basis of our knowledge of their knowledge. Aside from this operational definition of conceptual knowledge, the cognitive structure of "conceptual understanding" or a "concept" can be hard to define [12,24]. One theory holds that concepts are constructed at the moment from smaller cognitive structures or *resources* that are activated based on the perceived needs of the current task [25–27]. Mechanisms that can, to some extent, predict which resources will be activated and therefore how students will likely respond to a given task are being investigated, and this investigation is aided by dual-process theories of reasoning [9,11,12,28,29].

But dual process theories also offer an intriguing insight into what might be considered "conceptual understanding," especially if such understanding is operationally measured as performance on a set of physics tasks. Speaking to this point, Heckler [12] states, "If, as is suggested in […] the work on dual systems, the performance on these tasks is inevitably influenced by unconscious, automatic, bottom-up processes, then our understanding of understanding a science concept must include both explicit reasoning and automatic, bottom-up processes. One might say that both "System 1" and "System 2" are a necessary part of what we operationally mean by understanding a science concept, as they both may influence performance on any task relevant to the science concept. Indeed, a significant portion of expert science knowledge may be implicit." [12].

In other words, conceptual understanding is a product of an entangled with implicit, automatic thinking just as much as it is entangled with explicit or reflective thinking. One cannot become an expert, in this view, by only attending to reflective processes. Instead, one must also gain productive reflexive thinking. However, even if reflexive processes could be disentangled from conceptual understanding, and merely contribute to the cueing, framing, or transfer of conceptual knowledge to a given task, attending to reflexive thinking in pedagogy will reduce errors stemming from domain general processes. To improve performance on questions probing conceptual understanding, then, we must teach reflexive processes in addition to teaching reflective processes.

Heckler and Scaife investigated this premise with two methods of training related to answering physics questions about graphs where the quantities are related via a derivative (such as velocity being the derivative of position with respect to time) [13]. In the rule-training group, students were informed of the rule that the physical quantity could be found by looking at the slope of the graph and then gave two example problems with feedback. This aims at reflective processes as it explicitly confronts the student with the concept. In the example-training group, students attempted to learn, by trial and error, the correct response to a series of graph tasks, but the rule was not explicitly given. This type of training could be seen as targeted toward reflexive learning processes. As a control, there were a group of students that received no training at all. Both training groups evidenced better performance than the control group on a post-test assessment. In other words, both treatment groups overcame the previously observed fluency-based bias. Moreover, it was found that those in the example-based training shifted in their response time—the time to process slope decreased to the same time it took to process height—whereas, in the rule-based training, no shift in response time was observed. This suggested that the two groups overcame the fluency-based bias through different means—the example-training group reduced processing time and therefore reduced processing time errors, while the rule-training condition apparently allowed the students to consider the alternate dimension of slope through executive control.

The researchers conclude "the response times provide evidence that the different kinds of instruction can affect explicit and implicit mechanisms in different ways. Better knowledge of these mechanisms may help us to design instruction to improve student understanding of science concepts." [13]. The current manuscript is in large part a response to this call: we seek to explore the impact of instruction that attends to reflexive, type 1 processing in comparison to instruction that targets the reflective, type 2 processes.

Utilizing reflexive approaches that target the automatic type-1 reasoning processes in learning has been explored outside of physics education. As one example, perceptual learning modules that target type 1 processes were used in a mathematics context to improve student fluency with extracting structure from mathematical representations (i.e., graphs and equations) and performance on subsequent tasks [30]. Similar to these perceptual learning modules, Rau [31] designed activities to improve student fluency to make connections across visual representations in chemistry (e.g., Lewis structure, Bohr model, and electrostatic potential map). Related research inside physics education is also beginning to be done. In a physics context, DPTOR was used as a framework for developing instructional materials that produced improved performance on buoyancy questions [10]. Elby produced a series of tutorials that purposefully elicited a process 1 response that was inconsistent with formal physics and then guided students through a type 2 reasoning process that acknowledged and reconciled the process 1 response with formal physics reasoning [32]. One of these tutorials was demonstrated to have a bigger impact than tutorials that did not explicitly attend to type 1 processes in their construction [33].

In this study, we chose a specific context that has a rich conceptual landscape and also has wide importance in introductory physics: Newton's second law as it relates to friction. We will, therefore, explain some of the recent research related to friction and Newton's second law. Friction is a good context to study because there are two standard pathways for determining a value for the strength of a friction force—through inference via Newton's second law, and/or by direct calculation using the formula $f = \mu N$—and the surface features of a given problem can influence which of the two pathways students pursue. As an example, student reasoning has been investigated in the context of friction using a paired question methodology, in which a screening question—which students tend to answer correctly with correct reasoning—is followed by a target question that targets the same conceptual knowledge and utilizes similar lines of reasoning but is more difficult for students. This paired question methodology was used to study a static friction task in which students are expected to reason with Newton's second law to determine the magnitude of a friction force on a box that remains at rest [34].

In the screening question [see Fig. 1(a)], a single box at rest is shown with an applied force of 30 N acting on it. Students are asked to compare the magnitude of the friction force to the magnitude of the applied force. To answer correctly, one can reason that since the box is at rest, Newton's second law implies that the horizontal forces (friction and applied) must sum to zero and therefore the magnitudes of the two forces must be equal to each other. Approximately, 83% of students answered the screening question correctly [34].

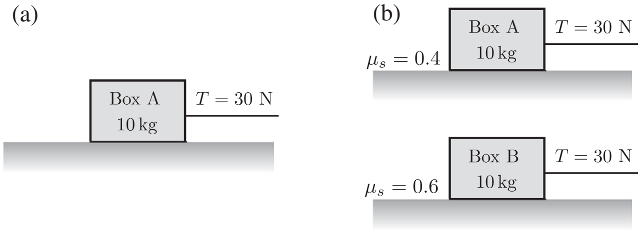

FIG. 1. Diagrams given to students for (a) the screening question and (b) the target questions of the two-box friction task.

In the target question [see Fig. 1(b)], two separate, identical boxes at rest on different surfaces are shown with identical applied forces of 30 N. The coefficient of static friction for each box-surface pair is shown next to each box. One could again reason that since the boxes are at rest, Newton's second law implies that the horizontal forces on each box (friction and applied) must sum to zero and therefore the friction forces on both boxes are equal to the $30 =$ N applied force. Only 65% of students answered the target question correctly, however, with 35% of students opting instead to focus their reasoning on the coefficients. To those students, the magnitude of the friction force on box A must be less than the magnitude of the friction force on box B because the coefficient for box A is less than the coefficient for box B.

Intriguingly, more than 20% of those students who answered the screening question correctly (with correct explanations) used the coefficient-based reasoning on the target question. Despite the fact that these students demonstrated the ability to assemble conceptual knowledge into a coherent line of reasoning on the screening question, they abruptly abandoned that line of reasoning on the target question. (In one interview excerpt, this abrupt shift was observed real time.)

These results were interpreted through a dual-process lens as a demonstration of students' reliance on reflexive process 1 thinking, in this case, cued by the distracting feature (the coefficients). The researchers suggested that an intervention related to metacognition might help to cause students to engage their reflective process 2 and see that building knowledge around the coefficients was less productive to this problem, but this proved unsuccessful. However, a modest intervention centered on directly challenging student satisfaction with the coefficient model was successful at improving performance on the target question—but only for those students who answered the screening question correctly with correct explanations [34].

Here, we use this screening and target question pair as part of a pre- and post-test to analyze the effect of two methods of instruction: one that promotes reflexive processes and one that promotes reflective processes. We aim to directly compare the two approaches to determine whether reflexive, bottom-up approaches to learning can be useful in producing equivalent or better performance shifts compared to more typical reflective, top-down approaches in the context of friction. Our study addresses the following two research questions:

> *Research question 1 (RQ1): Does training focused on bottom-up, reflexive reasoning processes produce equivalent or better performance shifts as reflective top-down approaches?*
>
> *Research question 2 (RQ2): How does a classroom intervention aimed at reflexive reasoning processes impact student qualitative inferential reasoning compared to a more reflective top-down approach?*

III. METHODOLOGY

The experiment was conducted via a Qualtrics survey given to calculus-based introductory physics 1 courses at two universities. All relevant instruction (using current research–based pedagogies) had been given at the time the survey was given, and the tasks were used as nongraded exam reviews. We collected data from two institutions during the fall of 2019. Institution 1 is a large public research university in the northeastern part of the United States, while institution 2 is a mid-size regional public university in the Pacific Northwest. At Institution 1, data were collected in one large-enrollment course with a single instructor. All students took the same laboratory, and students completed the survey prior to a midsemester exam for participation credit. Institution 1 had a 60% participation rate. At Institution 2, we collected data for the same course across six different instructors. All students took the same laboratory, which used research-based pedagogy. Students completed the survey in the last week of the quarter as a review exercise for final participation credit. Institution 2 had an 83% participation rate.

A three-question pretest was administered to students, following which the student was randomly assigned to receive either reflective or reflexive training. The reflective approach was meant to model a typical activity that might be given to students as a way to get them to reflect on the nature of Newton's second law in the context of static friction. The reflexive approach was meant to draw attention away from the distracting feature (the coefficient of friction) and toward the relevant feature (the forward force). Immediately after completing the training tasks, students were again given the three questions they saw on the pretest. The students did not know beforehand that there would be a post-test. It was noted at the beginning of the post-test, however, that they would be seeing the same questions that they had previously answered and that they were being given a chance to revisit them.

A. Pre- and Post-test

The pre- and post-test consisted of three questions: the screening and target question pair from Ref. [34] separated by an additional, far transfer question. The additional question is shown in Fig. 2.

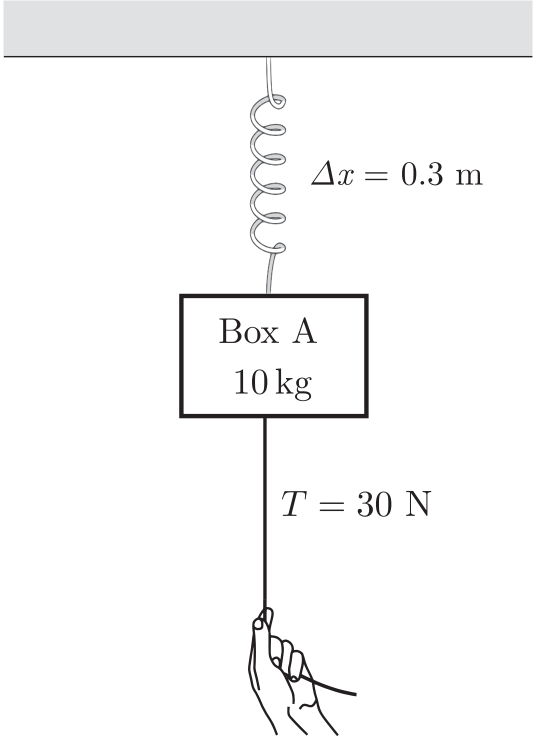

FIG. 2. The second of three questions in the pre- and post-test for this experiment. The first question posed on the test is the screening question shown in Fig. 1(a) and the third question is the target question shown in Fig. 1(b).

In this question, a box is hanging at rest on a spring while a downward force is applied. Students were asked to determine the magnitude of the spring force, which can be accomplished by applying Newton's second law and summing the downward forces of gravity and tension. The spring displacement given (0.3 m) is not needed and is also not useful in determining the spring force without also knowing the spring constant. The correct answer would then be 130 N (if g is rounded to 10 N/kg). This question is useful in determining whether students in the reflexive training condition were trained to simply answer whatever the tension force was or whether they would be oriented toward Newton's second law approach to the problem.

B. Reflective approach explained

The students in the reflective approach condition were given three practice problems. Each problem had a box at rest with an applied force and a resisting force, (question 1) a spring, (question 2) a friction force, or (question 3) a contact force from a hand (see Fig. 3). The first question had extra variables (the spring constant and displacement) that could be used to obtain the correct answer but was not needed. The second question had an extra variable (the coefficient of static friction) that could be useful in

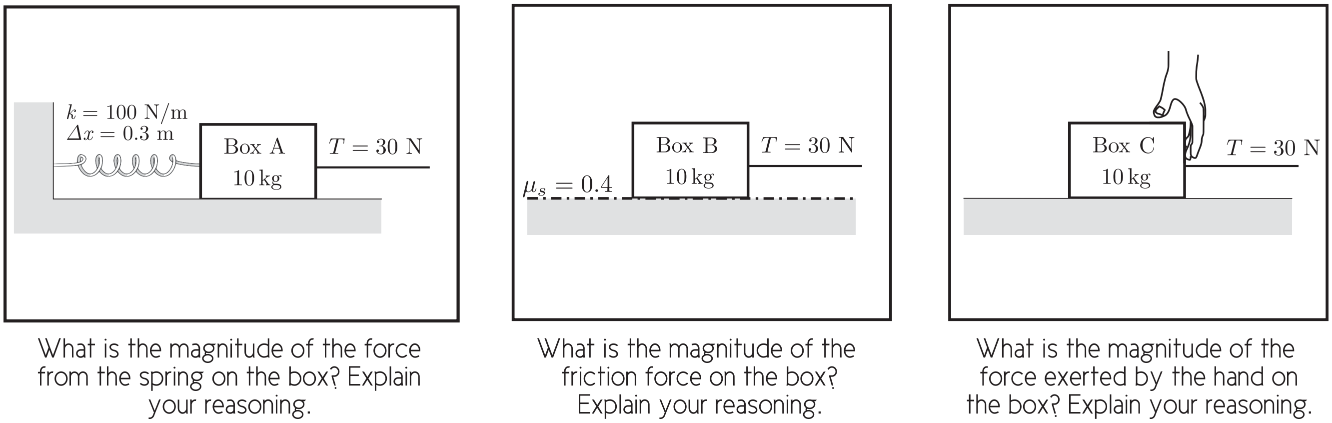

FIG. 3. The figures for the three tasks given in the reflective condition. In each case, the student was asked to find the magnitude of the force of friction, the spring force, or the force applied by the hand. After all three questions were answered, a follow-up question asked the student to consider why certain information was not needed to determine these forces.

sensemaking but is more commonly used to answer incorrectly. The third question had no extra variables, constraining the correct reasoning pathway.

Two follow-up questions were asked: "Was your reasoning on questions 1, 2, and 3 similar or different? Explain." and "In some of these experiments, information was given to you about the forces directed to the left (e.g., the spring constant). These values are not needed to solve the problem correctly. Explain why."

After these follow-up questions, the students in the reflective condition were given worked-out solutions to the three problems. In these explanations, students were told that while the spring constant and displacement *could* be used, they did not *need* to be used to solve the problem

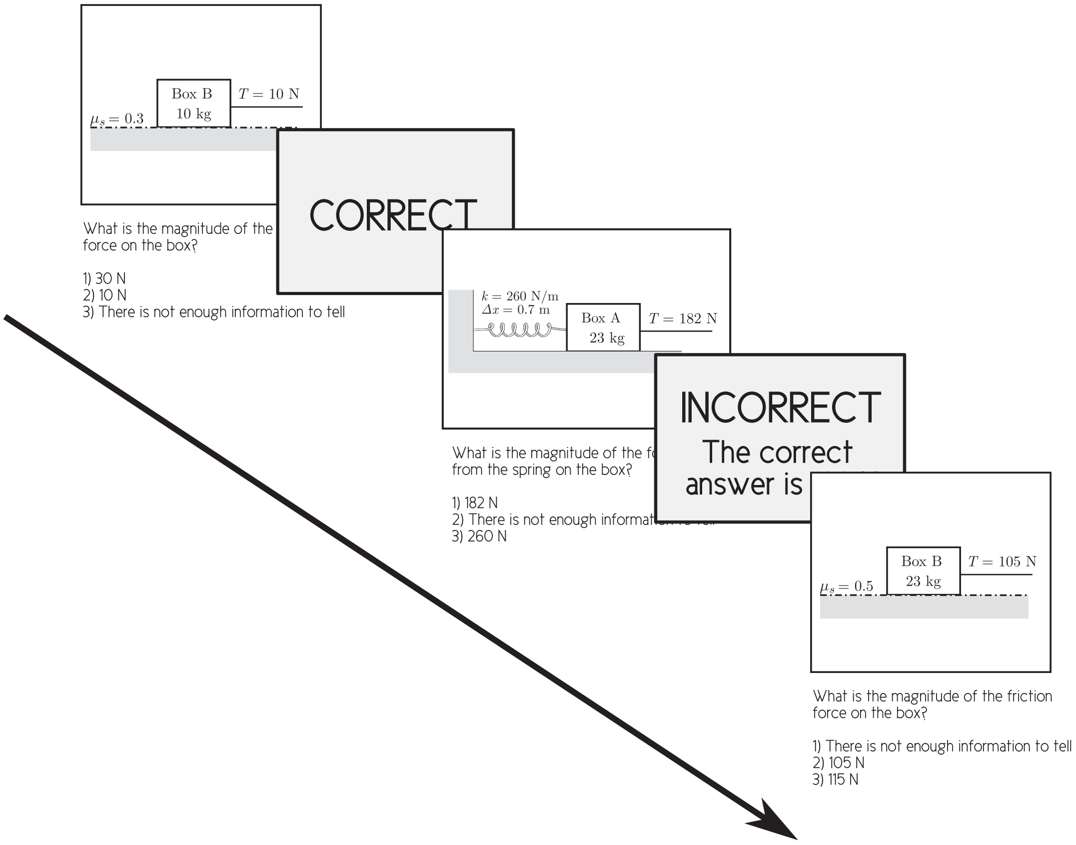

FIG. 4. Examples of the flow of questions posed in the reflexive training condition. Students received 10 of each of the three question types (Box A, B, and (C) in randomized order with the values of the numbers changed on each iteration.

and that while the coefficient of friction was useful in thinking about the *maximum* static friction between the surfaces, it would not yield the magnitude of that force in this context. Instead, students were told, they would need to reason from Newton's second law.

C. Reflexive approach explained

The students in the reflexive approach condition were asked to complete 30 multiple-choice problems "as quickly as you can while still remaining accurate." They were not asked to explain their reasoning and were provided with some feedback: after each question, the student was immediately told if they were correct or incorrect but with no further explanation.

Each student received the same set of questions. The 30 questions consisted of three question types (10 Q each) identical to the three used in the reflective condition with the exception that the values of the numbers were changed in each trial. An example of each type of question is shown in Fig. 4.

IV. RESULTS

To better understand the impact of each type of training, student responses were analyzed along a variety of different dimensions. Each of these is discussed in turn below. For the following analysis, our criteria for inclusion required that a student answer both Q1 and Q3 on the pre- and post-test. A student who did so, regardless of whether they answered Q2, was included in the analysis. However, we note that of those who answered Q1 and Q3 on both the pre- and post-test, only three students did not answer Q2 on both the pre- and post-test and one did not answer Q2 at all.

A. Performance

To investigate RQ1 ("Does the reflexive training produce equivalent or better performance shifts as the reflective training?"), student answer choices were coded as either correct or incorrect, with a summary of performance results shown in Table I. Student performance on the screening target pair (Q1 and Q3) was generally strong, with the target question showing slightly less strong performance. The vertical spring question, Q2, proved difficult for most students and saw no statistical improvement post-training.

At both institutions, the screening and target pair saw a statistical improvement in the reflexive condition with medium to large effect sizes using a chi-square test with Cramer's V. The reflective condition only saw a statistical difference at institution 2, and this was of smaller effect size compared to that of the reflexive condition at institution 2. A summary of the test statistics is given in Table II.

The responses to Q2 were analyzed with reference to the prevalence of "30 N" answers (i.e., the answer associated with the answer based on the tension force alone). Table III shows the percentage of 30 N responses for each category. While the prevalence of the 30 N answer increased after training, a 2×2 chi-square test determined that there was no statistical difference between the two conditions. For institution 1, $\chi^2 = 0.123$, $p = 0.73$, $V = 0.029$ and for institution 2, $\chi^2 = 0.0257$, $p = 0.87$, $V = 0.014$.

Students were able to complete the activity at their own pace in a location of their choosing. However, timing data reveal that students in either condition took roughly 45 min to complete the activity.

TABLE I. Percentage correct by question on the pre- and post-test, broken down by training condition, for each institution.

	Institution 1				Institution 2			
	Reflective (N: 81)		Reflexive (N: 95)		Reflective (N: 150)		Reflexive (N: 147)	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Q1	77.8	88.9	66.3	91.6	72.7	88.0	68.7	97.3
Q2	32.1	34.6	20.0	20.0	41.3	40.7	49.7	47.6
Q3	71.6	81.5	58.9	83.2	65.3	84.0	67.3	93.2

TABLE II. Results of 2×2 Chi-square test based on comparisons of correct and incorrect vs pre and post. The tests were performed for each condition at each institution. Bold-face p-value indicates statistical significance, and bold-face effect-size (V) indicates medium to large effect size.

	Institution 1								Institution 2							
	Reflective (N: 81)				Reflexive (N: 95)				Reflective (N: 150)				Reflexive (N: 147)			
	ChiSq	p value	V	df	Chi Sq	p value	V	df	Chi Sq	p value	V	df	ChiSq	p value	V	df
Q1	3.600	0.058	0.149	1	18.240	**0.000**	**0.310**	1	11.161	**0.001**	0.193	1	42.510	**0.000**	**0.380**	1
Q2	0.111	0.739	0.026	1	0.000	1.000	0.000	1	0.014	0.907	0.007	1	0.123	0.726	0.020	1
Q3	2.200	0.138	0.117	1	13.537	**0.000**	**0.267**	1	13.816	**0.000**	**0.215**	1	31.015	**0.000**	**0.325**	1

TABLE III. Percentage of responses to Q2 that answered 30 N.

	Institution 1		Institution 2	
	Pre	Post	Pre	Post
Reflexive (%)	18	62	20	47
Reflective (%)	16	48	18	45

B. Predifference and Postdifference scores—screening target pair

Since Q2 was difficult, showed no improvement post-training, and because our main interest is in the screening and target pair with a long history of analysis in the literature, we ignored Q2 in subsequent analysis. However, Q2 is useful in interpreting our data and will be referenced again in our discussion.

To gain more detail regarding the nature of the improvement in answering as a result of the two conditions, a difference score was computed by scoring the pretest and post-test separately out of 2 (1 point for each correct answer choice regardless of reasoning given) and taking the difference of the post-test score and the pretest score. Note that there are two ways a student could achieve a difference score of 1: (a) a student could answer both questions incorrectly on the pretest and one of the questions correctly on the post-test, or (b) a student could answer one question incorrectly on the pretest and answer both questions correctly on the post-test. The results for both institutions are shown in Fig. 5.

Because of low counts in some of the categories, Fisher's exact test was used to determine the significance of the apparent improvement in difference scores in the reflexive conditions. Fisher's exact test on the difference score for institution 1 (reflective vs reflexive) gives a p value of 0.038 with a medium effect size (Cramer's V) of 0.22. (This test was done on a 2×4 table of the data shown in Fig. 5.) An analysis of the residuals indicates that there are lower than expected counts in the reflexive condition for a difference score of -1 and higher than expected counts for a difference score of 2 (standardized residuals of -2.15 and $+1.77$, respectively). Fisher's exact test on the difference score for institution 2 gives a p value of 0.068 with a small effect size of 0.153; residuals show lower than expected counts in the reflexive condition for a difference score of 0 and higher than expected counts for a difference score of 2 (standardized residuals of -2.407 and $+1.98$, respectively).

Taken together, these results suggest a small effect showing that the reflexive condition produces a larger improvement, particularly for those who answer both questions incorrectly on the pretest.

C. Reasoning on target question postintervention

To answer RQ2 ("How does each training impact qualitative reasoning?"), the reasoning provided by students in justifying their answer choice for the target question (Q3) was categorized to assess for two main components: (1) whether they utilized reasoning that centered on Newton's second law or (2) whether their response focused only on the forward force without reference to Newton's second law. Criteria for the first category (N2L) included mentioning that the boxes were at rest or the acceleration was zero, that horizontal forces needed to be balanced, or that the net force was zero. The criteria for the second category (forward force) included mentioning the fact that the forward applied force was the same for both boxes but without reference to any other component of the second law reasoning line. An example response for each category is given below.

> *N2L: "The magnitude of friction acting on the boxes can only be equal to the external force acting on the object. Both forces are equal because both boxes remain at rest. This means that each friction force is the same".*
>
> *Forward Force: "Equal because the applied forces are the same, so the opposing frictional forces will also be the same".*

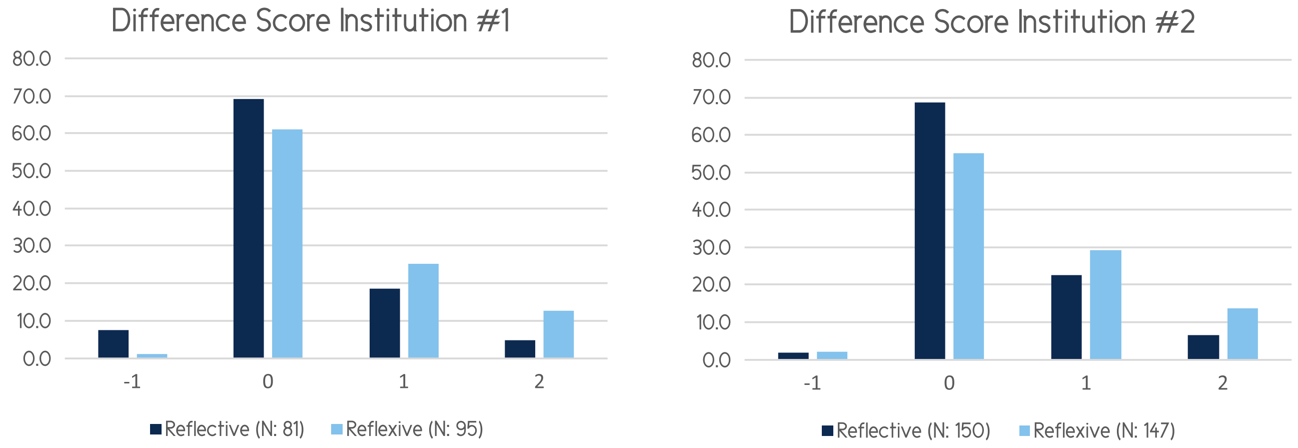

FIG. 5. Bar charts showing the percentage of students in each condition that achieved a specific difference score.

TABLE IV. Percentage of correct student responses to the target question (Q3) that fell into each reasoning category. The majority of correct responses were accompanied by correct reasoning.

	Institution 1		Institution 2	
Reasoning category	Reflective (*N*: 66)	Reflexive (*N*: 79)	Reflective (*N*: 125)[a]	Reflexive (*N*: 137)
Forward force (%)	15.2	11.4	8.8	11.7
N2L	77.3	63.3	84.8	79.6
No reasoning given	7.6	25.3	6.4	8.8

[a] At institution 2, one student's response was classified as "Other" and was excluded from this table.

TABLE V. Reasoning given on the target question (Q3) by those students who answered the target question incorrectly on the pretest but answered correctly in the post-test.

	Institution 1		Institution 2	
	Reflective (*N*: 12)	Reflexive (*N*: 25)	Reflective (*N*: 31)	Reflexive (*N*: 40)
Forward force (%)	17	20	10	20
N2L (%)	58	48	84	73
No reasoning given (%)	25	32	6	8

The categorization criteria were developed by reviewing student responses as well as wanting to assess the possible impact of the reflexive training—i.e., whether the reflexive training helped students to use Newton's second law (relying on the information that the box was at rest and therefore the forces were balanced) or whether the students simply learned that the forward force was always the right answer. Categorization was done by two of the authors and reviewed for consistency. Less than 2% of codes were found to be in conflict between coders, and these were resolved through discussion.

A summary of the results of this categorization scheme is given in Table IV. In all conditions—even in the reflexive condition—students who answer Q3 correctly appear to predominately use reasoning which references Newton's second law or uses it as a basis. For institution 1, a chi-square test shows a difference between the reflective and reflexive condition ($p = 0.04$, $V = 0.2$) and examination of the residuals shows that the "No Reasoning Given" category had higher than expected counts in the reflexive condition (+2.82 std residual). For institution 2, there was no difference between the conditions ($p = 0.54$, $V = 0.07$).

Likewise, we could analyze the students who answered the target question incorrectly on the pretest but who answered it correctly on the post-test. What reasoning did these students employ? Table V shows that students in this category predominately used Newton's second law reasoning in defense of their answers. There was no statistical difference found between the two conditions at either institution ($p = 0.90$, $V = 0.1$ for institution 1 and $p = 0.52$, $V = 0.15$ for institution 2).

V. DISCUSSION OF RESULTS

Question 1 is designed (and used elsewhere in literature) to probe the presence of (and ability to recall) basic conceptual knowledge of Newton's second law in the context of friction. Question 3 is designed to examine student ability to productively navigate this knowledge specifically when a salient distracting feature is present. Question 2 is designed for the current study to investigate student ability to use knowledge of Newton's second law in a novel and more complicated situation. At both institutions, students in the reflexive training condition had a statistical increase in performance on Q1 (the screening question) and Q3 (the target question) from pre to post with a higher effect size than the reflective training condition. Generally, this implies that the reflexive condition was more productive for students in answering those two questions. However, neither training was able to help students transfer knowledge about balancing forces to Q2, where there was more than one applied force.

Because we saw a larger positive shift in performance in Q1 in the reflexive condition compared to the reflective condition, we would conclude that, overall, the students in the reflexive condition were better able to construct knowledge of Newton's second law in the context of that problem. The larger positive shift in performance for Q2 shows that students in the reflexive condition were better suited to navigate their knowledge of Newton's Second Law in the presence of the salient distracting feature of the coefficients. However, the lack of statistical performance shift pre to post for Q2 demonstrates that neither training gave students the ability to navigate or construct their knowledge in a novel situation outside of friction. Thus, Q2, while largely disregarded in our analysis, still yields useful information for the purposes of this study.

It is not surprising that the reflexive training condition had an increase in performance on the screening and target questions. What is important, however, is that students largely used correct explanations in support of their answers. One might assume that the increase in performance was simply due to training students that the answer is always the value of the applied force without any reference to Newton's second law. This view would need to account for the fact that a comparable proportion of students in the reflexive condition used correct second law explanations in their responses to Q3 compared to the reflective training condition. Additionally, one could examine the response patterns on the second question. If students were trained to simply view the value of the applied force and answer with that value, then the answer 30 N (the value of the applied force in that problem) should become popular among the reflexive crowd and not among the reflective crowd. We examined this question and found that the relative prevalence of the 30 N answer was not statistically different between the two conditions on the post-test.

It may still be the case that students were trained to accept the value of the applied force without referencing any physics conceptual knowledge, and indeed the goal of the reflexive training condition was to focus attention on the applied force. One model of student reasoning on the target question assumes that students use the coefficients to build their reasoning because it is the first available default model [34]. In our study, students in the reflexive condition would be trained to have a new default model—the forward force—and they would then assemble correct reasoning to justify their answer based on this answer. We view this process as becoming more expertlike in constructing Newton's second law in problem solving when salient distracting features are present. However, we would need additional tests to determine if the reflexive training did indeed shift the default model in Q3, and it is clear from the results of Q2 that the more expertlike behavior is limited in scope.

At institution 1, the reflexive condition has a significantly higher number of students who gave no reasoning at all. One might interpret this as meaning that the reflexive condition trains students in the right answer but leaves them unable to articulate correct lines of reasoning that lead to that answer. However, further investigation of those 20 students revealed that 10 of them answered both questions in the screening and target pair correctly both times, and two more of these students answered the target question correctly in the pretest, suggesting that these students were not benefited from the intervention. (Perhaps these students, who presumably had robust knowledge of Newton's second law before the intervention, simply felt little need to explain their reasoning again following the 30 Q reflexive training.) Thus, if the reflexive training only "programed" the correct answer in the students without also helping them reason more effectively, it did so for a small enough number of students that it is not readily discernible compared to a more traditional, reflective approach.

It is notable that the reflexive training seemed to help most of those who were demonstrating the least conceptual knowledge on the pretest questions—that is, those students who answered both the screening and target questions incorrectly on the pretest. Could it be that those students, who had already received research-based instruction on Newton's second law, needed to gain some fluency with these types of questions in order to make sense of the conceptual knowledge they had presumably learned during the other course components? More research would be needed to determine the answer to this question.

These results bolster a claim tested elsewhere [9,11,34]. That is, after a training designed to draw attention away from a distracting feature and toward the relevant feature, students correctly assembled and used a line of reasoning consistent with correct conceptual knowledge. This result suggests that the distracting feature in the task (the coefficients) was preventing students from accessing or assembling the relevant, correct conceptual understanding, and, furthermore, once that barrier was removed, students were able to demonstrate correct reasoning on the problem. Thus, an incorrect answer to a physics question may not stem from incomplete conceptual understanding, but rather from processing difficulties unrelated to the physics content itself.

This has interesting implications for what we mean when we speak of conceptual knowledge. If the demonstration of conceptual knowledge is dependent on type 1 processes, then, to echo Heckler [12], our attempts to bolster conceptual knowledge must include attempts to affect type 1 processes with regard to physics questions. The key is to determine in which circumstances to target type 1 processes and which circumstances to target the reflective type 2 processes. In light of our results (and understanding that our intervention was done in addition to regular classroom instruction), we would conclude that a reflective approach to instruction on its own is apparently less efficacious when distracting features are present, and a better method of instruction, in this case, is one that includes interventions targeted at type 1 processes. This result would need to be tested in other conceptual contexts that also include distracting features.

In this study, we rely on the general design of the two types of training to differentiate between top-down, type-2 processing, and bottom-up, type-1 processing. We did not ascertain through experiment whether attention was really drawn away from the distracting features of the three tasks employed in the training (for instance, through eye tracking or other methods), though an increase in performance on the target question strongly suggests that attention was shifted away from the distracting feature. Instead, we opted to use an approach that would be more germane to physics instructors—to design activities using principles drawn from cognitive science and deploy these activities in a setting familiar to students. Further studies would need to be done to determine if the reflexive training was truly impacting type 1 processing in a meaningful way. However, the two designs (reflective and reflexive) are strikingly different, and a meaningful difference in outcome was achieved. More research ought to be done to determine how reflexive training can be utilized to impact far transfer performance (to perhaps remedy our null result in Q2), to determine in what contexts reflexive training can achieve greater results than reflective exercises, and furthermore, to explore how the two approaches can best be utilized in cooperation to give students the best chance of learning to use correct physics knowledge.

VI. CONCLUSIONS

In this study, we examined the effect of a classroom intervention designed to reroute conceptual reasoning pathways by directing attention away from a distracting feature. We found that the activity was successful in improving performance on target questions involving friction and that students who completed the activity utilized correct conceptual reasoning in a proportion comparable to those completing an activity with a reflective approach to instruction. This result, interpreted through a dual-process lens, gives us added insight into how to give attention to both type 1 and type 2 processes in instructional approaches. In the context of friction, an approach aligned with process 1 was just as good at improving performance on post-test questions, if not better, than an approach focused on process 2. It also seems apparent that the reflexive approach bolstered conceptual understanding (as measured by performance on the questions coupled with an analysis of their reported reasoning) among those who were previously unable to demonstrate correct reasoning on friction questions. More research is needed to know if this proves true in other physics contexts, or if in some contexts, a reflective approach leads to greater conceptual reasoning and learning.

ACKNOWLEDGMENTS

The authors wish to acknowledge valuable discussions with and contributions from Zachary Hutchinson and MacKenzie Stetzer at the University of Maine and Joshua Smestad and Joe Leahy III at the University of New England.

[1] J. S. B. T. Evans and K. E. Stanovich, Dual-process theories of higher cognition, Perspect. Psychol. Sci. **8**, 223 (2013).

[2] D. Kahneman, *Thinking, Fast and Slow* (Macmillan, New York, 2013).

[3] P. R. Heron, Testing alternative explanations for common responses to conceptual questions: An example in the context of center of mass, Phys. Rev. Phys. Educ. Res. **13**, 010131 (2017).

[4] L. C. McDermott and P. S. Shaffer, *Tutorials in Introductory Physics* (Pearson College Division, Upper Saddle River, NJ, 2001).

[5] M. C. Wittmann, R. N. Steinberg, and E. F. Redish, *Activity-Based Tutorials: Introductory Physics, The Physics Suite* (Wiley, New York, 2004).

[6] D. R. Sokoloff and R. K. Thornton, Using interactive lecture demonstrations to create an active learning environment, Phys. Teach. **35**, 340 (1997).

[7] D. Hestenes, Toward a modeling theory of physics instruction, Am. J. Phys. **55**, 440 (1987).

[8] E. Etkina and A. Van Heuvelen, Investigative science learning environment—A science process approach to learning physics, *Research Based Reform of University Physics* (AAPT, College Park, MD, 2007), https://psrc.aapt.org/items/detail.cfm?ID=4988.

[9] M. Kryjevskaia, M. R. Stetzer, and N. Grosz, Answer first: Applying the heuristic-analytic theory of reasoning to examine student intuitive thinking in the context of physics, Phys. Rev. ST Phys. Educ. Res. **10**, 020109 (2014).

[10] C. R. Gette, M. Kryjevskaia, M. R. Stetzer, and P. R. Heron, Probing student reasoning approaches through the lens of dual-process theories: A case study in buoyancy, Phys. Rev. Phys. Educ. Res. **14**, 010113 (2018).

[11] J. C. Speirs, M. R. Stetzer, B. A. Lindsey, and M. Kryjevskaia, Exploring and supporting student reasoning in physics by leveraging dual-process theories of reasoning and decision making, Phys. Rev. Phys. Educ. Res. **17**, 020137 (2021).

[12] A. F. Heckler, The role of automatic, bottom-up processes: In the ubiquitous patterns of incorrect answers to science questions, *Psychology of Learning and Motivation* (Elsevier, New York, 2011), pp. 227–267.

[13] A. F. Heckler and T. M. Scaife, Patterns of response times and response choices to science questions: The influence of relative processing time, Cogn. Sci. **39**, 496 (2014).

[14] L. C. McDermott, M. L. Rosenquist, and E. H. Zee, Student difficulties in connecting graphs and physics: Examples from kinematics, Am. J. Phys. **55**, 503 (1987).

[15] R. J. Beichner, Testing student interpretation of kinematics graphs, Am. J. Phys. **62**, 750 (1994).

[16] A. F. Heckler and A. M. Bogdan, Reasoning with alternative explanations in physics: The cognitive accessibility rule, Phys. Rev. Phys. Educ. Res. **14**, 010120 (2018).

[17] A. M. Bogdan and A. F. Heckler, Effects of belief bias on student reasoning from data, in *Proceedings of PER Conf., Portland, OR* (American Association of Physics Teachers, College Park, MD, 2013), 10.1119/perc.2013.pr.006.

[18] J. R. Sattizahn, D. J. Lyons, C. Kontra, S. M. Fischer, and S. L. Beilock, In physics education, perception matters, Mind Brain Educ. **9**, 164 (2015).

[19] J. C. Speirs, M. R. Stetzer, and B. A. Lindsey, Examining student tendency to explore alternate possibilities (to be published).

[20] R. Thornton and D. Sokoloff, Assessing student learning of Newton's laws: The force and motion conceptual evaluation and the evaluation of active learning laboratory and lecture curricula, Am. J. Phys. **66**, 338 (1998).

[21] D. Hestenes, M. Wells, and G. Swackhamer, Force concept inventory, Phys. Teacher **30**, 141 (1992).

[22] P. R. L. Heron, Empirical Investigations of Learning and Teaching, Part I: Examining and Interpreting Student Thinking, in *Proceedings of the International School of Physics "Enrico Fermi"*, Vol. 156 (Italian Physical Society, Bologna, Italy, 2004), pp. 341–350.

[23] L. C. McDermott, Oersted medal lecture 2001: Physics education research—The key to student learning, Am. J. Phys. **69**, 1127 (2001).

[24] A. A. DiSessa and B. L. Sherin, What changes in conceptual change?, Int. J. Sci. Educ. **20**, 1155 (1998).

[25] D. Hammer, A. Elby, R. Scherr, and E. F. Redish, *Transfer of Learning: Research and Perspectives*, edited by J. Mestre (Information Age Publishing Inc., 2005), pp. 89–119.

[26] D. Hammer, Student resources for learning introductory physics, Am. J. Phys. **68**, S52 (2000).

[27] E. F. Redish, A Theoretical Framework for Physics Education Research: Modeling Student Thinking, in *Proceedings of the International School of Physics "Enrico Fermi"*, Vol. 156 (Italian Physical Society, Bologna, Italy, 2004), pp. 1–63.

[28] A. Elby, What students learning of representations tells us about constructivism, J. Math. Behav. **19**, 481 (2000).

[29] M. Kryjevskaia, P. R. Heron, and A. F. Heckler, Intuitive or rational? Students and experts need to be both, Phys. Today **74**, 08, 28 (2021).

[30] P. J. Kellmann, C. M. Massey, and J. Y. Son, Perceptual learning modules in Mathematics: Enhancing students' pattern recognition, structure extraction, and fluency, Top Cogn Sci . **2**, 285 (2010).

[31] M. A. Rau, Making connections among multiple visual representations: how do sense-making skills and perceptual fluency relate to learning of chemistry knowledge?, Instr. Sci. **46**, 209 (2018).

[32] A. Elby, Helping physics students learn how to learn, Am. J. Phys. **69**, S54 (2001).

[33] T. I. Smith and M. C. Wittmann, Comparing three methods for teaching Newton's third law, Phys. Rev. ST Phys. Educ. Res. **3**, 020105 (2007).

[34] M. Kryjevskaia, M. R. Stetzer, and T. K. Le, Failure to engage: Examining the impact of metacognitive interventions on persistent intuitive reasoning approaches, in *Proceedings of PER Conf., Minneapolis, MN* (American Association of Physics Teachers, College Park, MD, 2014), 10.1119/perc.2014.pr.032.