Evaluative sensemaking: frequency of and variance among instructors

Travis Herring, MacKenzie Lenz, Paul J. Emigh, and Elizabeth Gire
Department of Physics, Oregon State University, Corvallis, OR 97331

Kelby T. Hahn
College of Education, Oregon State University, Corvallis, OR 97331

Evaluative sensemaking strategies are used by physics experts and students to reflect on their answers. To understand how students make meaning of and check the correctness of their answers, we identify various strategies students used on homework problems in three different sections of a calculus-based introductory physics course. We use an elaborative coding process to categorize students' responses to an explicit reflection prompt for circular motion, rotational motion, oscillations, and optics problems. The evaluative sensemaking strategies most commonly used were asserting the correctness of units, arguing reasonableness based on given problem parameters or the solution process, and drawing on prior knowledge from outside physics. The frequency of strategies varied by instructor and were aligned with instructional emphases. Students who experienced instruction that emphasized concepts tended to use more conceptual strategies, while students who experienced more procedural/algebraic instruction employed more mathematical strategies.

I. INTRODUCTION

Experts use sensemaking strategies to find solutions to unfamiliar problems and check that their solutions are appropriate [1, 2]. Their use of sensemaking strategies, such as considering conserved quantities and possible idealizations, frequently leads to a more fruitful process or correct answer. Evaluative sensemaking strategies used at the end of a problem, such as checking units and considering limiting cases, help experts build confidence in the reasonableness of their solutions. Experts use these evaluative sensemaking strategies when solving unfamiliar problems, and they implicitly expect students to use them as well [2].

Some targeted instructional approaches have helped students learn to do evaluative sensemaking, such as considering special cases, checking units, and judging the reasonableness of numeric answers [3, 4]. We are interested in characterizing the breadth of strategies students use when asked to evaluate their answers to physics problems. As a first step in understanding how to help students develop more expert-like sensemaking skills, we investigate how introductory physics students perform evaluative sensemaking on homework problems. We ask two main research questions:

1. What evaluative sensemaking strategies do introductory physics students use when prompted to reflect on their answers to problems and with what frequency?
2. Do students with different instructors use different evaluative sensemaking strategies?

II. INSTRUCTIONAL CONTEXT

In this study, we analyze student work from the second term of the three-term introductory physics with calculus sequence in Fall 2016 at Oregon State University—a large, public research institution. The introductory sequence consists of three 10-week terms. The second term of the introductory physics sequence covers a variety of topics: uniform circular motion and orbital motion, rotational energy, moments of inertia and torque, simple harmonic motion, driven and damped oscillations, sound and light waves, and finally wave interference, wave optics, and ray optics.

The course has 3 components: lecture, studio, and lab. Students attend two 50-minute lectures each week, where physics content is presented in a traditional lecture format, often interspersed with clicker questions. Students attend 2-hour studio sections once per week where they interact in small groups and work with the concepts learned in lecture. Studio is led by the students' instructor and 2 additional instructional team members. Students get hands-on experience with the physics content through weekly 2-hour labs. The second of these course components—studio—is why the course is colloquially referred to as the studio course, and will be referred to as such throughout the remainder of this paper.

There were 3 instructors referred to by the following pseudonyms: Empower Eric, In-my-head Isaac, and Concept Carl. Each instructor taught their own lecture and studio sections; student enrollment in lab sections was determined by instructor but taught by teaching assistants. Each section had different exam problems and in-class instruction, but the same homework problems, labs, and studio activities.

TABLE I. Evaluative sensemaking elements of required solution format and grading rubric. Students were given both the format and the rubric and required to address all elements of the rubric for full credit.

Required Solution Format: Reflect. Is the answer reasonable? Does it make physical sense? – Evaluate the result. Is the answer reasonable? Are the units correct? Does the answer make sense in limiting cases? Does the answer make physical sense? Include a written explanation for why the answer makes sense and what it implies about the physical system.
Grading Rubric: A clear and complete explanation is given for why the result makes sense (or does not make sense if the incorrect answer was reached), and what it tells us about the physics of the situation.

Interviews from previous research by author ML included the self-reported sensemaking pedagogy of each instructor [2]. Empower Eric's sensemaking pedagogy focused on student ability to check answers mathematically, believing this mathematical sensemaking gives students power over their own learning. Concept Carl's sensemaking pedagogy focused on connecting physics context to the math the students were doing. In-my-head Isaac's sensemaking pedagogy was between Empower Eric's and Concept Carl's in that they taught students sensemaking mathematically with some focus on contextual sensemaking.

III. METHODS

Six written homework assignments from weeks 1, 2, 4, 5, 6, and 8 of the 10-week studio class were collected and scanned by ML. Homework assignments consisted of two problems that students answered following the required solution format and TAs graded the homework according to the grading rubric (Table I). Students were given both the solution format and the rubric prior to completing the assignments. This solution format prompt was analyzed to understand students' use of evaluative sensemaking strategies.

Homework was collected from 113 students total: 51 from Empower Eric, 34 from In-my-head Isaac, and 28 from Concept Carl, though not every student completed every assignment. On average each assignment was turned in by 47 of Empower Eric's students, 27 of In-my-head Isaac's students, and 25 of Concept Carl's students, for a total average of 99 students' homework per assignment.

Data analysis started with a codebook generated by ML from a study of student work in the first term of the same introductory physics series, taught by the same 3 instructors in Spring 2016 [2]. Starting from ML's codebook, TH used a thematic analysis approach to perform elaborative coding on the homework from the second term of the series [5, 6].

Inter-Rater Reliability (IRR) tests were conducted to ensure integrity of the codebook [7]. During the coding process, ML and TH each coded 7 different homework assignments—three assignments from Empower Eric and two assignments each from In-my-head Isaac and Concept Carl. A similarity score was found for each IRR test by dividing the number of agreed upon codes for a given assignment by the total number of codes identified by each coder.

$$\text{Similarity Score } = \frac{\text{Total Codes } - \text{ Mismatched Codes}}{\text{Total Codes}} \times 100\%$$

TH and ML each coded 3 problems for Empower Eric with each problem's similarity score being $> 95\%$. Similarly, 2 different problems for each Concept Carl and In-my-head Isaac were coded with similarity scores also $> 95\%$. These high similarity scores represent strong agreement between the coders, indicating that the final codebook depicts an accurate representation of student use of sensemaking strategies.

IV. RESULTS AND DISCUSSION

A total of 2143 codes were applied over all 6 assignments for all 113 students. We describe the categories of the strategies students employed and their frequency. We then discuss differences in student strategy use among instructors.

A. *Evaluative Sensemaking Strategies and Frequency*

Our results, shown in Table II, illustrate the types of evaluative sensemaking strategies that students use and how often they use them. In this context, we calculated a percent frequency that refers to a code application per-problem per-student calculation. The higher the frequency percent, the more often a particular student used that code in their reflection for their solution to a given homework problem.

$$\text{Frequency } = \frac{\text{Total Code Applications}}{\text{12 HW Problems} \times \text{99 Students}} \times 100\%$$

The codes are grouped into categories that represent broader themes. We found 6 categories of strategies: 'Units', 'Initial Conditions', 'Properties of Answer', 'Procedural/Approach', and 'Other Strategies'. All justification by unit arguments are contained within the Units category. The Conceptual category contains strategies where students used conceptually-based arguments. The Initial Conditions category encompasses strategies where any initial conditions that the students have to work with prior to the reflection prompt appear. Analyzing particular aspects of solutions—such as magnitude, sign, and functional behavior—is categorized as Properties of Answer. Procedural/Approach encompasses strategies involving the particular path that a student uses to solve problems. The Other Strategies is a catch-all for strategies that do not fit within other categories.

The most common categories of sensemaking strategies that students used when reflecting on their answers were Units (55%), Conceptual (48%) and Initial Conditions (36%).

Within the six categories there were an array of strategies students employed. The most commonly employed strategy was *Mentions appropriate units* (46%), which makes up most of the Units category. *Mentions appropriate units* is considered straightforward reasoning—we are using the term 'straightforward reasoning' to refer to reasoning that uses given information inherent to the problem being solved (*i.e.*, what values are given, what units the answer must be in, what the problem context is, *etc*.). *Mentions appropriate units* is straightforward reasoning because students asserted that the units worked without any displayed work; students may not know if the units in their algebra are correct but can claim they are correct based on the units given in the problem statement.

The next most commonly used strategy was *Argument based on physical system* (30%), which makes up most of the Initial Conditions category. This strategy is considered straightforward reasoning because students use given information about the problem to reason about their solution. Students do not need to infer or assume things about the problem to use the *Argument based on physical system* strategy.

The third most common strategy is *Assumes correct reasoning or process* (15%) within the Procedural/Approach category. When students engaged in this strategy, they made arguments that assumed that the process they used to find their solution was correct. This is straightforward reasoning because students used their own generated work to reason about their solution.

Unsurprisingly, straightforward reasoning dominates in student strategy use as it depends on the context given by problem statements. More sophisticated reasoning—using information not stated in the context of the problem or student-generated work—was less common.

B. *Differences Between Instructors*

Knowing how instructors may influence student sensemaking gives insight about how students respond to different sensemaking pedagogies. If instructors do influence the types of sensemaking that their students use, then instruction can be tailored to desired student sensemaking outcomes.

We conclude that the overall sensemaking behavior of the students in a course aligns with the self-described sensemaking pedagogies of their respective instructors. However, we acknowledge that there are a limited number of sensemaking pedagogies (mathematical and conceptual) presented in this data set. Because of this limited exposure to pedagogy, students are limited to the types of strategies they would frequently be exposed to.

Empower Eric's sensemaking pedagogy was mathematically motivated—his focus was on answer checking in an algebraic way such that students gain agency over their an-

TABLE II. Codebook and occurance percent of evaluative sensemaking strategies between 3 introductory physics sections. The percentage shown is the frequency of a given code occurring per homework problem per student.

Category / *Code*	Excerpt from student work	Freq
Units		**55%**
Mentions appropriate units	"the units work as expected..."	*46%*
Shows an explicit unit check	"Part a) $\frac{m}{s^2} = \frac{rev}{min} \times \frac{min}{sec} \times \frac{m}{rev} \times \frac{1}{sec} = \frac{m}{s^2}$ ..."	*9%*
Conceptual		**48%**
References non-physics prior knowledge	"hawks are large and can fly pretty fast..."	*14%*
Compares like values	"the top force is much smaller than the bottom..."	*10%*
References prior physics knowledge	"acceleration is the vector sum of tangential acceleration and radial acceleration."	*8%*
References physical past experience	"in the real world pendulum oscillations take place relatively quickly..."	*6%*
Makes an estimation	"its less than $\frac{1}{2}$ the period $T = 2.0$ sec, so its within the 0 to 2π rad range."	*4%*
Compares unlike values	"the total tension needs to be greater than the mass of the beam."	*4%*
Interprets physical meaning	"we learn that the pendulum swings to 23° on either side of the vertical..."	*2%*
Initial Conditions		**36%**
Argument based on physical system	"0.689 s seems like a reasonable amount of time for a pendulum to move..."	*30%*
References assumptions/laws/concepts	"Given the sound waves travel equally in all directions... the formula aligns..."	*6%*
Properties of Answer		**18%**
Mentions magnitude or range of number	"The answers are in the correct approximate range."	*7%*
"Reasonable" with no other description	"answer is reasonable"	*6%*
Mentions sign of number	"The answer should be positive given the bird is moving toward the person."	*3%*
Interprets functional behavior	"The squared s' term guarantees two locations that the image will be focused."	*1%*
Mentions extremes or limiting case	"the angle difference is equal to θ_p, then θ_i must be getting very close to 90°."	*1%*
Procedural/Approach		**17%**
Assumes correct reasoning or process	"the calculations were right and the formulas used were correct..."	*15%*
Assumes correct use of math	"if all the math is right then the units should be right."	*2%*
Other Strategies		**10%**
Uncommon novice sensemaking	"The value of V was correct, but the values for h were not."	*4%*
Uses external source	"The problem makes sense... because Mastering Physics says so."	*4%*
Re-states solution only	"output is 5.03 W if the sound intensity level is 90 dB at a distance of 20 m."	*2%*

swer. The most frequent sensemaking strategies used by their students were *Mentions appropriate units* (55%), *Argument based on physical system (39%), Shows an explicit unit check* (14%), and *Uses external source* (13%). These strategies are mathematically based and coincide with Empower Eric's sensemaking pedagogy.

Concept Carl's sensemaking pedagogy is conceptually motivated—his focus was on using conceptual arguments to reason about math, linking the physics to the algebra. The most frequent strategies used by their students were *Mentions appropriate units* (38%), *Argument based on physical system* (25%), *References non-physics prior knowledge* (24%), and *Assumes correct reasoning or process* (23%). These strategies are conceptually based and coincide with Concept Carl's sensemaking pedagogy.

In-my-head Isaac's sensemaking pedagogy was both mathematically and conceptually motivated—he stressed the connection of physics concepts to the mathematical operations but also primarily used sensemaking in an answer checking manner. This pedagogy lies somewhere between Empower Eric and Concept Carl. The most frequent strategies used by their students were *Mentions appropriate units* (41%), *Argument based on physical system* (21%), *References non-physics prior knowledge* (19%), and *Assumes correct reasoning or process* (13%). These strategies are both mathematically and conceptually based and coincide with In-my-head Isaac's sensemaking pedagogy.

V. IMPLICATIONS AND CONCLUSIONS

Students evaluated their answers to homework problems in a wide variety of ways. Some of these strategies are productive at detecting possibly incorrect answers (like explicitly

TABLE III. The most common sensemaking strategies for students from each of the three different instructors alongside each instructor's sensemaking pedagogy. The listed strategies are the four most frequent strategies used by each instructor's students. However, the two most frequent strategies were the same for each instructor while the third and fourth were only the same for In-my-head Isaac and Concept Carl. Frequencies were calculated in the same manner as in Table II, except the average number of students for each assignment (46, 28, and 25 students respectively for Empower Eric, In-my-head Isaac, and Concept Carl) for each respective instructor was used (*i.e.,* frequency is the number of code applications per homework problem per average number of students for a given instructor).

Instructor ***Code***	**Sensemaking Pedagogy**	**Freq**
Empower Eric	Math-based approach: unit checking and reasoning about problem parameters.	
Mentions appropriate units		*55%*
Argument based on physical system		*39%*
Shows an explicit unit check[1]		*14%*
Uses external source[1]		*13%*
In-my-head Isaac	Balanced approach: unit checking and relating past experiences.	
Mentions appropriate units		*41%*
Argument based on physical system		*21%*
References non-physics prior knowledge[2]		*19%*
Assumes correct reasoning or process[2]		*13%*
Concept Carl	Conceptual approach: conceptual reasoning when solving example problems.	
Mentions appropriate units		*38%*
Argument based on physical system		*25%*
References non-physics prior knowledge[2]		*24%*
Assumes correct reasoning or process[2]		*23%*

The first and second most common strategy used by students were the same for all instructors.
[1] Empower Eric's students used more mathematical sensemaking strategies as their 3rd and 4th most common.
[2] In-my-head Isaac's and Concept Carl's students used the same, more conceptual sensemaking strategies as their 3rd and 4th most common.

checking units). The breadth of strategies suggests the collection of students have many different resources for evaluative sensemaking.

Infrequently, students used some physically unsound strategies, like *Compares unlike values* (4%). Other strategies suggest a very superficial process, like *Uses external source* (4%) and *"Reasonable" with no other description* (6%). Such strategies were used infrequently by students and did not receive full credit. These unsound or superficial strategies are unlikely to strengthen students' understanding of the physics problem they are solving.

The most common strategies observed were similar across the three instructors: *Mentions appropriate units* (46%) and *Argument based on physical system* (30%). We consider both of these strategies to be straightforward reasoning. Like the potentially unproductive strategies above, the common strategies are relatively easy to implement but can provide students with more confidence that their answers are correct. However, even though these two strategies were common, students' work was unfortunately often superficial in the same manner of the unproductive strategies above. This may be an effect of student pragmatism; sophisticated strategies that require more skill or a wider knowledge base may also require more time and resources (intuitions, experiences, or formal physics knowledge) than are available.

We identified two types of sensemaking pedagogies (mathematical and conceptual) employed to different degrees by the three instructors. The limited number of pedagogies suggests that students may not learn new sensemaking strategies. In any given situation, many sensemaking strategies can be productive. First, we suggest that instructors use consistent, formal names for sensemaking strategies so that students know what they are and can begin learning when it is appropriate to use them. Second, instructors should support students in practicing a variety of specific sensemaking strategies to increase students' proficiency in performing them. Third, instructors should discuss how students can identify productive strategies that leverage their skills, intuitions, and prior formal physics knowledge.

ACKNOWLEDGMENTS

We would like to thank the OSUPER group and the students and instructors who participated in this study.

[1] C. Singh, When physical intuition fails, Am. J. Phys. **70** (2002).
[2] M. Lenzand E. Gire, Faculty views of and expectations for dimensional analysis, PERC Proceedings (2016).
[3] T. Sikorski, G. White, and J. Landay, Uptake of solution checks by undergraduate physics students, in *Physics Education Research Conference 2017*, PER Conference (Cincinnati, OH, 2017) pp. 368–371.
[4] A. Warren, Impact of teaching students to use evaluation strategies, Phys. Rev. ST Phys. Educ. Res. **6** (2010).
[5] V. Braunand V. Clarke, Using thematic analysis in psychology, Qualitative Research in Psychology **3** (2006).
[6] J. Saldaña, *The Coding Manual for Qualitative Researchers* (SAGE Publishing, 2013).
[7] H. E. A. Tinsleyand D. J. Weiss, Interrater reliability and agreement of subjective judgments, Journal of Counseling Psychology **22**, 358 (1975).