Examining the Relationship of Scientific Reasoning with Physics Problem Solving

Carol Fabby and Kathleen Koenig
University of Cincinnati

Abstract

Recent research suggests students with more formal reasoning patterns are more proficient learners. However, little research has been done to establish a relationship between scientific reasoning and problem solving abilities by novices. In this exploratory study, we compared scientific reasoning abilities of students enrolled in a college level introductory physics course for their ability to solve problems requiring different levels of conceptual understanding. Reasoning abilities were measured by Lawson's Classroom Test of Scientific Reasoning and students' ability to correctly solve different levels of problem on examinations was measured by the Taxonomy of Introductory Physics Problems. Results indicate that students with higher reasoning abilities perform equally well across problem levels while students of lower reasoning abilities struggle in solving problems that depend on higher conceptual understanding. This suggests that students with lower reasoning abilities may depend more readily on basic recall of facts and simple procedures to solve problems.

Keywords: scientific reasoning, problem solving, conceptual learning, undergraduate

Introduction

As the nation prepares students for future science careers by incorporating the Next Generation Science Standards (2013) into K-12 curriculum, post-secondary institutions are working to increase the number of students graduating with degrees in the science, technology, engineering, and math (STEM) disciplines. Educators are undertaking several approaches to attract and retain students in STEM, including understanding how students in physics use their reasoning abilities to engage in problem solving. Understanding this relationship will provide educators with a foundation to enhance student knowledge through skill awareness and development. Many physics problems are complex and require students to integrate content knowledge with critical thinking to determine what the problem is asking so they may determine the best approach to resolving it. However, many students enrolled in introductory physics courses are novice problem solvers and may memorize problem types or may simply apply a set of solutions to problems with similar surface features.

As students progress in their first year of college physics, they may move along the continuum from novice to advanced problem solvers which may be in response to students transitioning from high school to college with various levels of reasoning abilities. Complex problem types require a higher level of reasoning which may not be fully developed in students entering college. If students cannot move beyond basic recall they may decide to change majors out of the STEM fields. The authors of this paper were interested to learn if a correlation exists between scientific reasoning abilities and problem solving performance on simple to complex physics problems. Therefore, it is important to understand how reasoning abilities relate to student ability to solve problems requiring different levels of critical thinking. Understanding how students engage in problem solving may provide insight into developing support mechanisms which may include various types of tutoring strategies, implementing assessments which combine reasoning and problem solving abilities, or redesigning lectures or laboratory settings to incorporate targeted reasoning skills.

This paper discusses how students of various reasoning abilities enrolled in an introductory physics course solved different types of problems based on the levels of critical thinking referred to as information and mental procedures by Teodorescu, Bennhold, Feldman and Medsker (2013). The taxonomy used to categorize the physics problems and assess to what cognitive level students' progress during problem solving is detailed below. A discussion of how this exploratory study compared the categorical assessment of problem solving for students with higher, average and lower reasoning abilities concludes the article.

Prior Literature

According to Inhelder and Piaget (1958) and Shrager and Siegler (1998), students naturally develop reasoning abilities as they progress through the stages of learning. Inhelder & Piaget (1958) indicated that although basic reasoning skills begin around age 4 in the preoperational stage (ages 4-7); scientific reasoning abilities typically develop in adolescence. At this stage individuals begin to include more complex logic in their thinking and consider multiple variables in problem solving, understand physical and social phenomenon, and consider the perspective of "if/then" reasoning.

Shrager and Siegler (1998), basing their research on Inhelder and Piaget's work, determined that individuals develop reasoning abilities at different rates and at different ages as indicated in their strategy choice model. Simple problem-solving strategies can be learned at earlier ages, while more complex strategies may develop with practice and maturation as the individual grows into adulthood. Shrager and Siegler indicate that one of the most advanced problem-solving strategies involves retrieval, a category they describe as the ability of an individual to integrate the knowledge and problem-solving skills he/she has previously learned into memory. As one grows older these retrieval skills are readily implemented through decision making and as adaptive problem-solving strategies, rather than depending on memorized procedures to be successful.

In more recent years, researchers have compared scientific reasoning abilities of students in introductory physics courses with gains in *conceptual* learning (Coletta & Phillips, 2010). This research suggests students with more formal reasoning patterns are more proficient learners. However, little has been done to investigate whether or not scientific reasoning abilities relate to one's ability to solve different cognitive levels of physics problems, including those based on retrieval of facts, performing simple calculations, integrating basic physics knowledge and those which require alternative representations through symbolization.

Understanding Physics Problem Solving Skills

One aspect of problem solving, problem categorization, was established by Chi, Feltovich and Glaser (1981) as they studied how expert and novice problem solvers categorized and represented a variety of physics problems. The results of their study have been used as a foundation for research to understand how students in introductory physics courses problem solve. Chi et al.'s research showed experts categorize problems into types defined by the major physics principles used in a solution and novices categorize them into types as defined by the entities contained in the problem statement. This means the categories constructed by the novices may not correspond to existing internalized understanding of the problem, but rather denote a brief representation of the surface features. Although both experts and novices use similar features in the problem categorization, experts base their selection on procedural knowledge and applicability; whereas novices base their determinations on information provided by the problem. This information on problem solving has been utilized by many researchers, including work by Tuminaro and Redish (2007) who established several categories based on cognitive processes, such as mapping meaning to mathematics, pictorial analysis, and transliteration to mathematics.

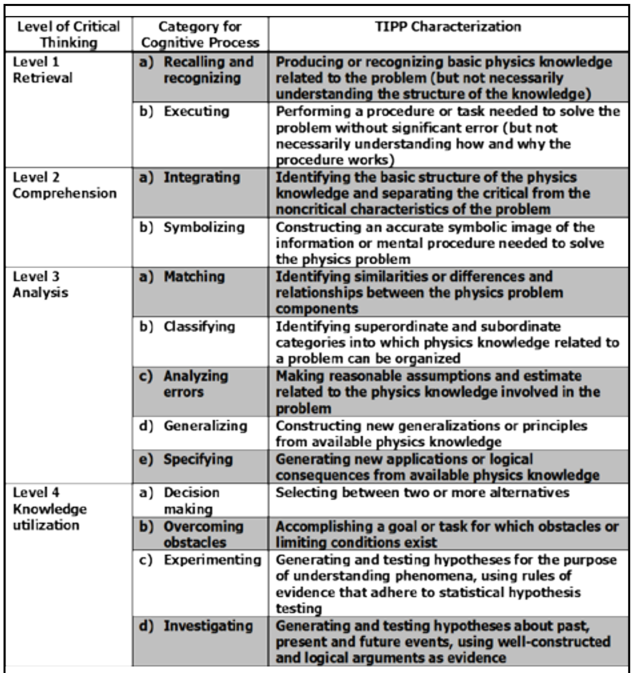
Taxonomy of Introductory Physics Problems

The Taxonomy of Introductory Physics Problems (TIPP) (Teodorescu et al., 2013) is a new framework useful in analyzing physics problems and understanding physics problem solving. It is based on three critical thinking domains (information, mental procedures and psychomotor procedures) which define the cognitive processes needed to solve physics problems. This physics-specific taxonomy has a foundation in the New Taxonomy of Educational Objectives (Marzano& Kendall, 2007) as it: "addresses problem solving, involves both knowledge domains and cognitive processes that have been identified by [Physics Education Research] PER as relevant for physics problem solving, and makes a clear distinction between the cognitive process and the knowledge involved in problem solving" (Teodorescu et al., 2013). The currently published taxonomy includes four cognitive levels of problem solving: retrieval, comprehension, analysis, and knowledge utilization. Each of these levels are described in terms of cognitive processes and procedures, which are based on previous research in problem solving, such as Tuminaro and Reddish's (2007) categories of mapping and data representation, and provide other researchers with a framework with which to construct physics problems to assess student problem solving abilities. By comparing student performance based on this taxonomy with reasoning abilities measured at the beginning of the course, establishing the existence of a relationship between reasoning and problem solving abilities was realized.

Methods

Population and Setting

This study was conducted at a large midwestern university. Data was collected from students enrolled in the first semester introductory physics course for students majoring in the health sciences during Fall Semester, 2013. There were 365 students enrolled in the course across three lecture sections taught by two professors. The course involved three hours of lecture, traditional in nature with modest use of a personal response system, where students use a computer or hand-held electronic device to respond to questions posed in lecture class. Of the 365 students enrolled in the course, 172 students were randomly selected across the three lecture sections to complete the Lawson Classroom Test of Scientific Reasoning (LCTSR) (Lawson, 1978, 2000). Of this subset, 133 students also completed all of the course exams.



Level of Critical Thinking	Category for Cognitive Process	TIPP Characterization
Level 1 Retrieval	a) Recalling and recognizing	Producing or recognizing basic physics knowledge related to the problem (but not necessarily understanding the structure of the knowledge)
	b) Executing	Performing a procedure or task needed to solve the problem without significant error (but not necessarily understanding how and why the procedure works)
Level 2 Comprehension	a) Integrating	Identifying the basic structure of the physics knowledge and separating the critical from the noncritical characteristics of the problem
	b) Symbolizing	Constructing an accurate symbolic image of the information or mental procedure needed to solve the physics problem
Level 3 Analysis	a) Matching	Identifying similarities or differences and relationships between the physics problem components
	b) Classifying	Identifying superordinate and subordinate categories into which physics knowledge related to a problem can be organized
	c) Analyzing errors	Making reasonable assumptions and estimate related to the physics knowledge involved in the problem
	d) Generalizing	Constructing new generalizations or principles from available physics knowledge
	e) Specifying	Generating new applications or logical consequences from available physics knowledge
Level 4 Knowledge utilization	a) Decision making	Selecting between two or more alternatives
	b) Overcoming obstacles	Accomplishing a goal or task for which obstacles or limiting conditions exist
	c) Experimenting	Generating and testing hypotheses for the purpose of understanding phenomena, using rules of evidence that adhere to statistical hypothesis testing
	d) Investigating	Generating and testing hypotheses about past, present and future events, using well-constructed and logical arguments as evidence

Table 1. Critical thinking domains, categories and characteristics used to describe problem solving. Table reproduced from (Teodorescu et al., 2013, pp. 3).

Tests and Scoring

Lawson Classroom Test of Scientific Reasoning (LCTSR). The LCTSR is a 24-question multiple choice test with questions in six reasoning domains: conservation of mass and volume, proportional thinking, identification and control of variables, probabilistic thinking, correlational thinking and hypothetico-deductive reasoning. This test has become a standardized tool for assessing student scientific reasoning abilities. The test consists of 11 paired questions in which students are asked to respond to a question and subsequently choose the best reason for their response, along with two independent questions. Students must answer both questions in each pair correctly in order to receive a point toward their final score. Therefore, the maximum score for the LCTSR is 13.

Course Examinations

The lecture course exams were designed such that all students in all sections took the same written test at designated times during the term. The exams consisted of three to four free response problems and five multiple choice problems; all similar to those found at the end of the chapters in the text. The free response problems required students to determine the correct concept and equation(s) necessary to perform relevant calculations, and make determinations regarding the solutions, some of which included reasoning for their problem solving choices. The multiple choice questions included problems which were either conceptually driven or quantitative in nature.

Ranking

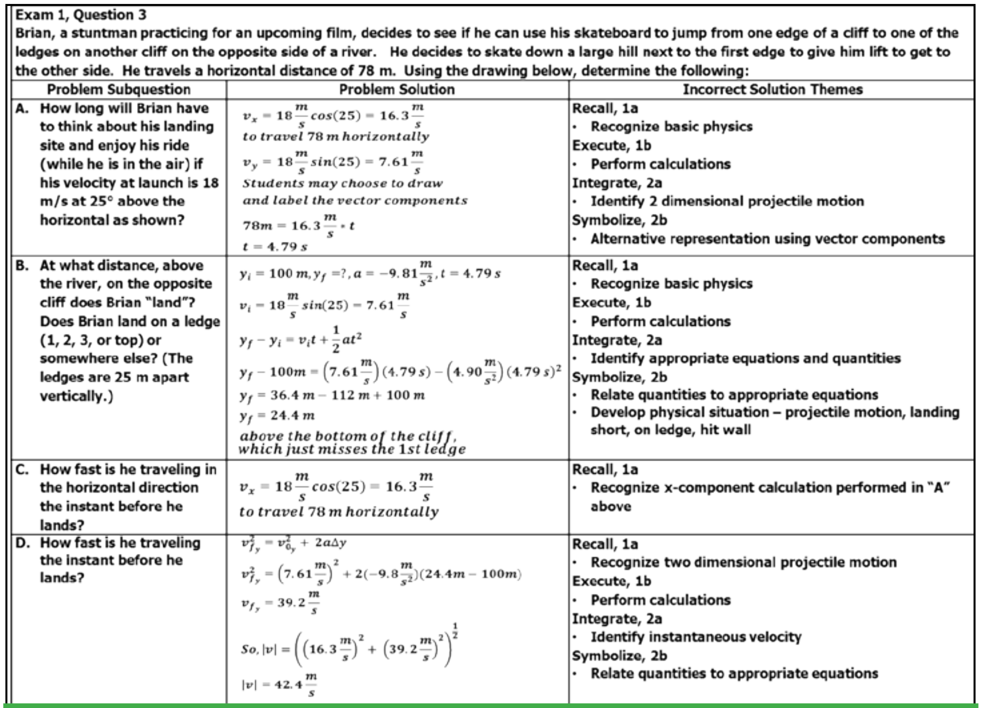
The 133 students in this study were split into three groups based on the LCTSR score distribution. The distribution of scores did not provide a natural split into three equally-sized groups. Therefore, we chose to define the groups such that the number of scores in the higher and lower reasoning groups would be approximately the same. There were 27 scores (20% of the sample) from students in the higher reasoning group (LCTSR scores of 10 to 13), 74 scores (56% of the sample) in the average reasoning group (scores of 6 to 9), and 32 scores (24% of the sample) in the lower reasoning group (scores of 1 to 5).

Problem Characteristics

The information and mental procedures as described in TIPP (Teodorescu et al., 2013) were used to categorize the targeted physics problems for the research detailed in this article. The terms *information* and *mental procedures* are defined in TIPP as a categorization of the knowledge and thought processes required to solve physics problems. Processes located higher on the taxonomy scale require a higher level of cognitive thinking to accomplish the task. A description of each critical thinking domain is included in Table 1.

Targeted Problems for Examinations

Just as reasoning abilities vary between students, so do problem solving abilities. To determine a relationship between the two abilities it is important to understand at what level students problem solve. Problems included on physics exams in a traditionally taught introductory physics course are typically derived from standard textbook questions. These problems are used to assess student understanding of basic concepts and applications of these concepts to physical situations. However, most traditional physics courses do not assess what cognitive processes are used in problem solving or how students progress along a continuum of problem solving.



Exam 1, Question 3
Brian, a stuntman practicing for an upcoming film, decides to see if he can use his skateboard to jump from one edge of a cliff to one of the ledges on another cliff on the opposite side of a river. He decides to skate down a large hill next to the first edge to give him lift to get to the other side. He travels a horizontal distance of 78 m. Using the drawing below, determine the following:

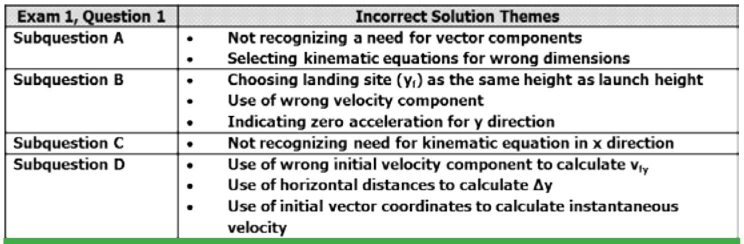
Problem Subquestion	Problem Solution	Incorrect Solution Themes
A. How long will Brian have to think about his landing site and enjoy his ride (while he is in the air) if his velocity at launch is 18 m/s at 25° above the horizontal as shown?	$v_x = 18\frac{m}{s}cos(25) = 16.3\frac{m}{s}$ *to travel 78 m horizontally* $v_y = 18\frac{m}{s}sin(25) = 7.61\frac{m}{s}$ *Students may choose to draw and label the vector components* $78m = 16.3\frac{m}{s} * t$ $t = 4.79\ s$	Recall, 1a • Recognize basic physics Execute, 1b • Perform calculations Integrate, 2a • Identify 2 dimensional projectile motion Symbolize, 2b • Alternative representation using vector components
B. At what distance, above the river, on the opposite cliff does Brian "land"? Does Brian land on a ledge (1, 2, 3, or top) or somewhere else? (The ledges are 25 m apart vertically.)	$y_i = 100\ m, y_f = ?, a = -9.81\frac{m}{s^2}, t = 4.79\ s$ $v_i = 18\frac{m}{s}sin(25) = 7.61\frac{m}{s}$ $y_f - y_i = v_i t + \frac{1}{2}at^2$ $y_f - 100m = \left(7.61\frac{m}{s}\right)(4.79\ s) - \left(4.90\frac{m}{s^2}\right)(4.79\ s)^2$ $y_f = 36.4\ m - 112\ m + 100\ m$ $y_f = 24.4\ m$ *above the bottom of the cliff, which just misses the 1st ledge*	Recall, 1a • Recognize basic physics Execute, 1b • Perform calculations Integrate, 2a • Identify appropriate equations and quantities Symbolize, 2b • Relate quantities to appropriate equations • Develop physical situation – projectile motion, landing short, on ledge, hit wall
C. How fast is he traveling in the horizontal direction the instant before he lands?	$v_x = 18\frac{m}{s}cos(25) = 16.3\frac{m}{s}$ *to travel 78 m horizontally*	Recall, 1a • Recognize x-component calculation performed in "A" above
D. How fast is he traveling the instant before he lands?	$v_{f_y}^2 = v_{0_y}^2 + 2a\Delta y$ $v_{f_y}^2 = \left(7.61\frac{m}{s}\right)^2 + 2(-9.8\frac{m}{s^2})(24.4m - 100m)$ $v_{f_y} = 39.2\frac{m}{s}$ $So, \lvert v\rvert = \left(\left(16.3\frac{m}{s}\right)^2 + \left(39.2\frac{m}{s}\right)^2\right)^{\frac{1}{2}}$ $\lvert v\rvert = 42.4\frac{m}{s}$	Recall, 1a • Recognize two dimensional projectile motion Execute, 1b • Perform calculations Integrate, 2a • Identify instantaneous velocity Symbolize, 2b • Relate quantities to appropriate equations

Table 2. This example illustrates the solution to the projectile motion problem and categories assigned by the authors to each subquestion.

As part of this research, the authors used the hierarchical structure in TIPP (Teodorescu et al., 2013) to assess to which cognitive level students progress during problem solving of three standard textbook problems on two exams. These problems included different aspects of Newton's laws (projectile motion, forces, acceleration, and tension) and were chosen because they were also based on the four categories in the first two levels of TIPP: retrieval of facts, performing simple calculations, integrating basic physics knowledge and those which require alternative representations through symbolization. For this research, none of the problems assessed went beyond level two of TIPP.

Initially, the authors of this paper individually categorized the three problems to determine the type of critical thinking and highest level of cognitive processes necessary to correctly solve the entire problem. Inter-rater reliability was established by comparing the two sets of categorizations and resolving minor discrepancies. A similar evaluation was performed to establish the highest problem solving level for the individual subquestions. The content was first identified and each aspect of how that portion of the problem should be solved was listed, mapping each to corresponding characteristics on the taxonomy.

To understand the subquestion categorizations, an example is provided in Table 2. For subquestions 'a', 'b', and 'd' students had to recognize basic physics and perform calculations which correspond to categories 'Recall: 1a', 'Execute: 1b' and 'Integrate: 2a'; while subquestion 'c' was categorized at level '1a', since students must recall this value from subquestion 'a'. Students must perform different processes to obtain the 'Symbolize: 2b' level in this example. Subquestion 'a' requires students to present an alternative representation of initial conditions by using vector components. Students must draw a vector diagram labeling the components or write out each vector component equation and correctly use the component values in calculations to achieve level '2b'. For subquestion 'b', students must be able to develop the situation presented in the problem where the skateboarder's projectile motion will end by hitting the opposing cliff face just short of the first ledge; while subquestion 'd' requires students to identify and relate quantities to appropriate equations to find the solution.



Exam 1, Question 1	Incorrect Solution Themes
Subquestion A	• Not recognizing a need for vector components • Selecting kinematic equations for wrong dimensions
Subquestion B	• Choosing landing site (y_f) as the same height as launch height • Use of wrong velocity component • Indicating zero acceleration for y direction
Subquestion C	• Not recognizing need for kinematic equation in x direction
Subquestion D	• Use of wrong initial velocity component to calculate v_{f_y} • Use of horizontal distances to calculate Δy • Use of initial vector coordinates to calculate instantaneous velocity

Table 3. Common themes identified from incorrect solutions for each subquestion from question one on exam one.

Process to Assess Student Problem Solving Abilities. Each student response to the designated exam problem was assessed using categories similar to those found in Table 2. The students' written responses on each exam subquestion were evaluated and the corresponding characteristics identified. From the list of characteristics each level of problem solving for that student's response on the corresponding subquestion was established. To determine the highest level of problem solving ability, each response must include a logical path toward the final solution. If a student demonstrates correct use of procedures and explanations then the highest categorization will be applied. However, if a response presents a disjointed solution by including a higher cognitive process without the corresponding lower processes, the highest categorization must be adjusted. This type of response may indicate the student has simply memorized solution parts or applied quantities to inappropriate equations rather than a full understanding of the steps required to solve the problem correctly, which requires a lower categorization. For the purposes of this research, if a student response was completely divergent of the physics concept or was left blank, a category of zero ('0') was designated. Errors propagated from a previous subquestion were not counted against the student's problem solving ability.

When evaluating incorrect student responses to establish the proper level of understanding, the authors looked for emerging themes in problem solving difficulties (See Table 3). Three examples of incorrect student responses are provided in Figs. 1-3 to demonstrate how categorization levels were distinguished. In Fig. 1 a common theme for subquestion 'a', an inability to acknowledge vector components when calculating the skateboarder's time of flight, is represented. This pattern indicates these students recognize the basics of projectile motion and have the ability to perform calculations, but may not fully understand how to approach solving the problem. This example shows the student understood the need for vector components, but does not use them in the calculation for time. Instead, he uses the initial velocity without consideration of the launch angle. Because there is a disconnect between the symbolic representation of information and the values used in the calculation, the category was reduced to a '1b'.

In subquestion 'b' a theme emerged indicating students were not able to develop the scenario by relating appropriate quantities to the proper equation or correctly

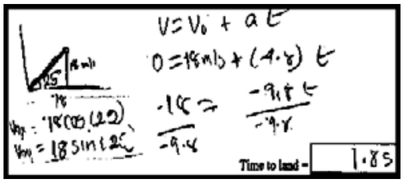
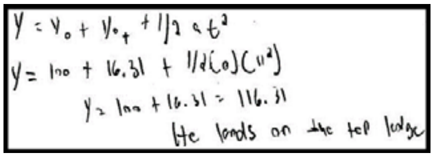


Figure 1. Example of an incorrect student response where vector components are not used.

determining where the skateboarder would land requiring a categorization of '1b'. This pattern demonstrated two aspects of misunderstanding: 1) students continued to use inappropriately selected kinematic equations to calculate at what point along the opposing cliff face the skateboarder landed, and 2) students indicated the landing site was at a similar level as the launch height by either using the horizontal component for the initial velocity or using a zero value for the acceleration in the y-direction. The example in Fig. 2 showcases both of the aspects of (2) above.



$y = y_0 + y_{0t} + 1/2\,a t^2$

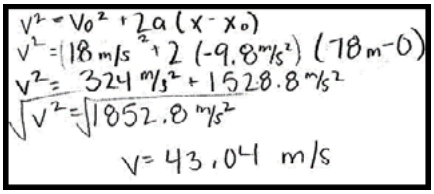
$y = 100 + 16.31 + 1/2(0)(11^2)$

$y = 100 + 16.31 = 116.31$

He lands on the top ledge

Figure 2. Example of a student response indicating use of incorrect vector component and incorrectly using a value of zero for the vertical acceleration.

A similar theme continues for subquestion 'd', where students were unable to identify the target concept of instantaneous velocity and did not adequately demonstrate construction of a mental procedure to correctly solve the problem. Several students indicated the velocity just before landing was the same as the initial velocity because there are no horizontal forces acting on the system, which netted a category of '1a' due to not recognizing two dimensions need to be considered for this projectile motion problem. Many other students struggled to identify the correct kinematic equations to use in determining the vertical velocity component and the instantaneous velocity. Typically these students calculated the final velocity with the correct equation but continued to use the initial velocity value without considering vector components. They used the correct equations to obtain the instantaneous velocity, but used the horizontal values for initial velocity and distance when calculating the vertical velocity component or a combination. The latter combination response is depicted in Fig. 3.



$v^2 = v_0^2 + 2a(x - x_0)$

$v^2 = (18\,m/s)^2 + 2(-9.8\,m/s^2)(78\,m - 0)$

$v^2 = 324\,m/s^2 + 1528.8\,m/s^2$

$\sqrt{v^2} = \sqrt{1852.8\,m/s^2}$

$v = 43.04\ m/s$

Figure 3. Example of a student response indicating no use of vector components and incorrect use of x-component values for distance.

Results and Discussion

The students' reasoning abilities at the beginning of the introductory physics course were determined by ranking their scores on the LCTR into three categories: higher reasoning ability, average reasoning ability and lower reasoning ability. Of the 133 students who completed the reasoning test, 20% were in the higher reasoning group (n=27; mean score=11 out of 13 or 85%), 56% in the average reasoning group (n=74; mean score=7 out of 13 or 54%), and 24% in the lower reasoning group (n=32; mean score=4 out of 13 or 31%).

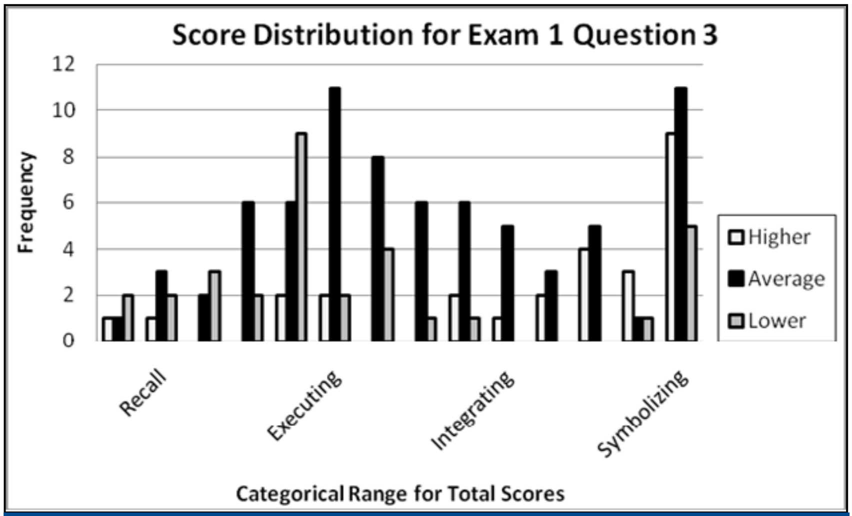


Figure 4. The overall student categorical performance on exam one, question three, compares the three reasoning groups by observing the clusters and peaks of score frequencies. The performance of the higher group increases toward its peak at Symbolizing; average reasoners peaked at Executing but are clustered between Executing and Integrating; and, the lower group is peaked just below Executing.

To determine if a relationship between reasoning abilities and problem solving exists, the problem solving performances on three exam problems were compared with reasoning scores, both within and between reasoning groups. Overall, a pattern emerged indicating students in the higher reasoning group outperformed those in the average and lower reasoning groups.

Overall Performance

The overall performance between reasoning groups on the three problems was assessed by comparing the problem solving score distributions. The following figures indicate the frequency of each category level by reasoning groups for the three targeted exam questions. Figure 4 indicates the majority of students in the higher reasoning group were successful in solving question three on exam one which corresponds to the highest level of cognitive process in this study (33% at the category level of Symbolizing and 37% at level Integrating; p=0.02 comparing higher to average and p=0.00 comparing higher to lower) than students in the average and lower reasoning groups (42% at Executing level and 47% at just below

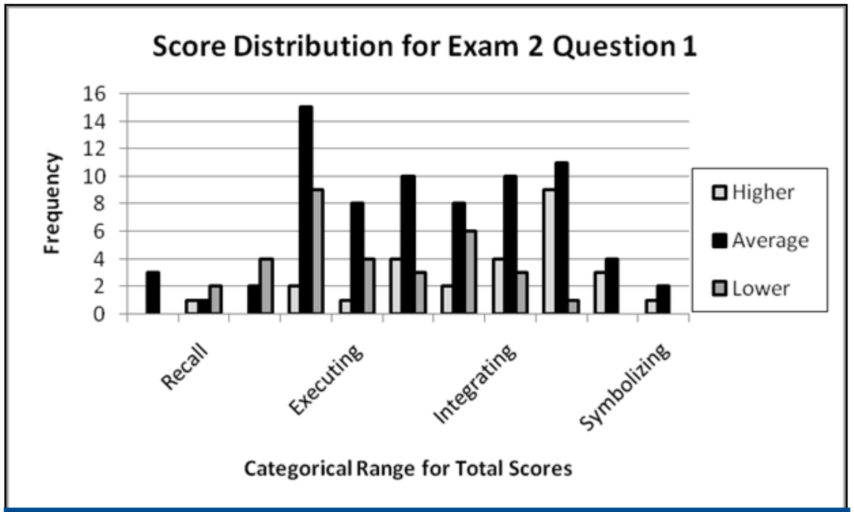
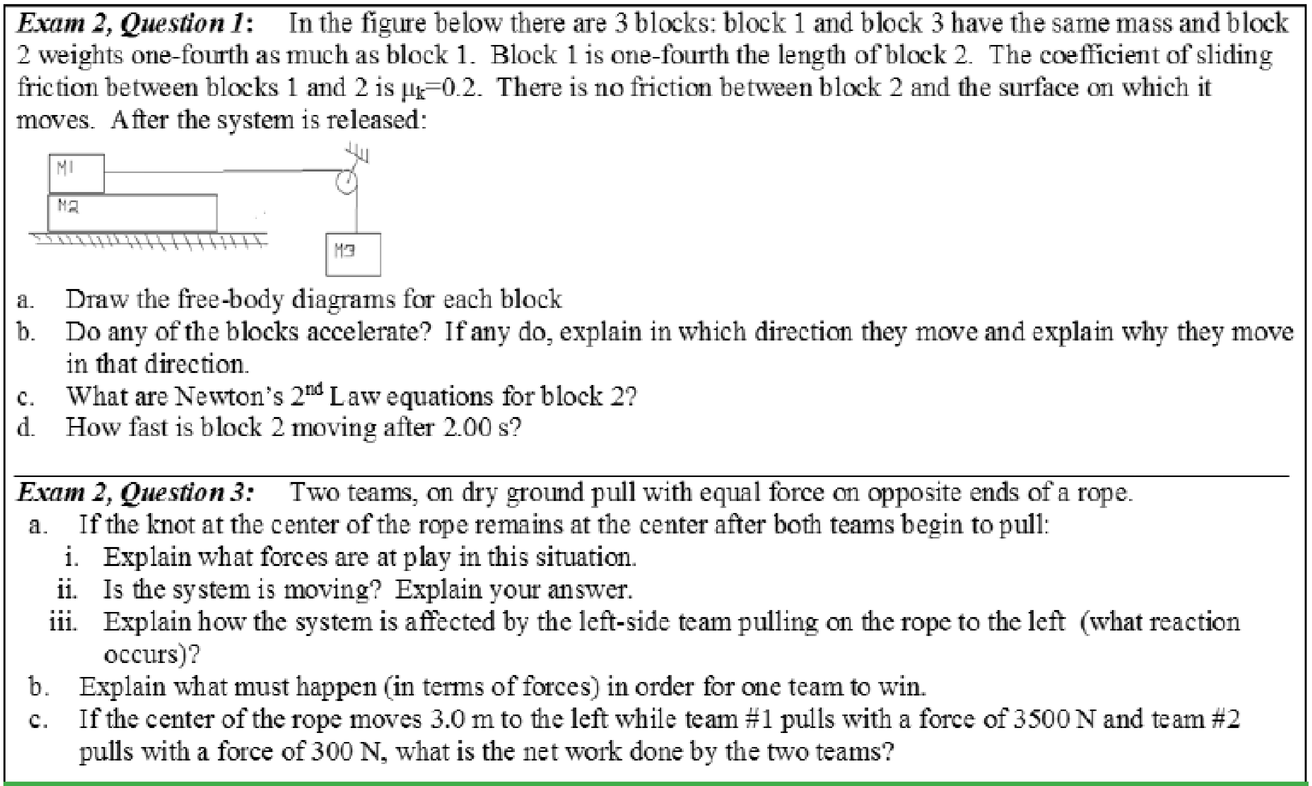


Figure 5. The overall student categorical performance on exam two, question one, compares the three reasoning groups. The data observed in this graph indicates similar performance by all three reasoning groups and are clustered in the center of the categorical scale. The higher reasoners' scores tend toward Integrating and the lower reasoners' scores are located just below Executing.



Exam 2, Question 1: In the figure below there are 3 blocks: block 1 and block 3 have the same mass and block 2 weights one-fourth as much as block 1. Block 1 is one-fourth the length of block 2. The coefficient of sliding friction between blocks 1 and 2 is $\mu_k=0.2$. There is no friction between block 2 and the surface on which it moves. After the system is released:

a. Draw the free-body diagrams for each block
b. Do any of the blocks accelerate? If any do, explain in which direction they move and explain why they move in that direction.
c. What are Newton's 2nd Law equations for block 2?
d. How fast is block 2 moving after 2.00 s?

Exam 2, Question 3: Two teams, on dry ground pull with equal force on opposite ends of a rope.
a. If the knot at the center of the rope remains at the center after both teams begin to pull:
  i. Explain what forces are at play in this situation.
  ii. Is the system is moving? Explain your answer.
  iii. Explain how the system is affected by the left-side team pulling on the rope to the left (what reaction occurs)?
b. Explain what must happen (in terms of forces) in order for one team to win.
c. If the center of the rope moves 3.0 m to the left while team #1 pulls with a force of 3500 N and team #2 pulls with a force of 300 N, what is the net work done by the two teams?

Table 4. The two additional problems used on course exams to assess problem solving abilities.

the Executing level, respectively; p=0.18 comparing average to lower). This means the higher reasoning group was successful in solving the problem which corresponds to the highest level of cognitive process determined for the problem. Students in the average and lower reasoning groups were not fully successful in solving the problem, but showed work that demonstrated lower levels of reasoning which corresponds to the lower problem solving categories.

Figure 5 indicates student performance is more similar between groups for question one on exam two. Table 4 provides details of this problem and question three on exam two, the remaining two problems assessed as part of this research. However, the overall performance of the higher reasoning group remains above the performances of the average and lower reasoning groups (48% between category levels of Integrating and Symbolizing, 38% between Executing and Integrating levels, and 53% between Recall and Executing levels, respectively; p=0.13 comparing higher and average, p=0.02 for higher and lower, and p=0.07 for average and lower).

Although there is little difference in student performance on question three on exam two (See Table 4), as presented in Fig. 6, the higher reasoning group continued to outperformed the average and lower reasoning groups (52% just above the Integrating category level, 64% at the level of Integrating, and 72% at level Integrating, respectively; p=0.53 comparing higher and average, p=0.18 for higher and lower and p=0.43 for average and lower).

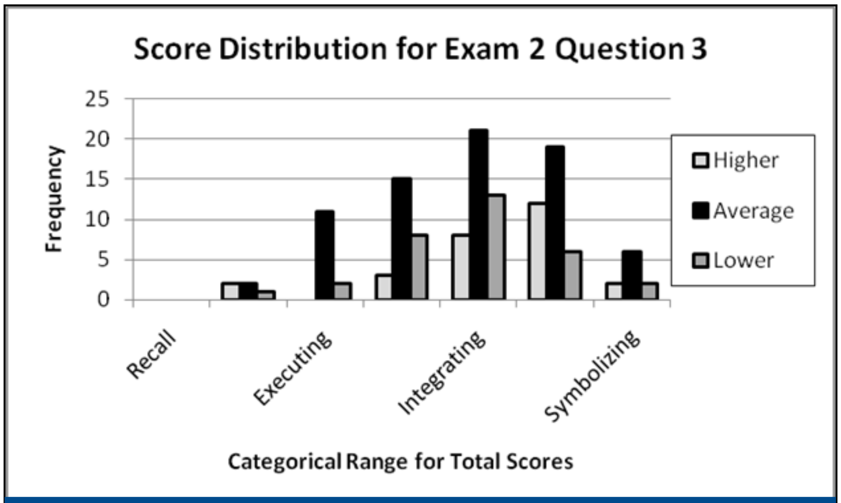


Figure 6. The overall student categorical performance on exam two, question three, compares the three reasoning groups. This graph indicates similar performance by all three reasoning groups in that few students performed to the highest categorical level. For those who did, their scores were clustered around the Integrating level.

Comparison of Problem Subquestions

The subquestions of the problems included on the exams were compared within and between reasoning groups to determine if any tasks showed differences in ability. Table 5 includes the subquestion data for the averages of each reasoning group. For this research, if an average of the categories placed the value between levels, it was rounded up to the next cognitive process, since this may indicate an increase in ability above the lower level. The maximum category levels correspond to the categories described in Table 2.

This data indicates the higher reasoning group outperformed the lower reasoning group in half of the problem subquestions and in three of the ten subquestions when compared with the average reasoning group. All groups reached the highest problem solving level for five categories on this scale (Exam 1, Question 3.c, Exam 2, Question 1.a and Exam 2, Question 3.b(i, ii, iv), 3.c and 3.e). However, the higher reasoning group also performed to the highest cognitive processing level on exam one, question 3.a and the second highest level on exam 1, question 3.b and 3.d and exam 2, question 1.c and 1.d&e. This suggests the higher reasoning group consistently completed the cognitive processing tasks at a higher level than the average and lower reasoning groups for these types of physics problems.

	Exam 1, Q3				Exam 2, Q1			Exam 2, Q3		
Subquestion label	a	b	c	d	a	c	d & e	b.(i,ii,iv)	c	e
Maximum Category Level Possible	2b	2b	1a	2b	2b	2b	2b	2b	1b	1b
	Average Category Level Achieved For Each Subquestion									
Lower Reasoning	1a	1b	1a	1b	2b	1b	1b	2a	1b	1b
Average Reasoning	1a	2a	1a	1b	2b	2a	1b	2a	1b	1b
Higher Reasoning	2b	2a	1a	2a	2b	2a	2a	2a	1b	1b

Table 5. Summary of average problem solving performance on problem subquestions by reasoning group. A higher level of achievement was observed in the higher reasoning group on Exam 1, Q3 and Exam 2, Q1. The overall average for Exam 2, Q3 was similar for all reasoners.

Projectile motion, exam one, question three. The pattern shown in Fig. 7 demonstrates a significant difference in problem solving abilities of the higher reasoning group on each of the subcategories as compared with average and lower reasoners. There were 74% higher reasoning students who achieved the maximum level on subquestion a, as compared with 45% average reasoners (p=0.05) and 19% lower reasoners (p=0.00). For subquestion b: 52% higher reasoners reached the maximum level as compared with 22% average (p=0.01) and 19% lower (p=0.00). Similarly for subquestions c and d: 89% higher reasoners performed well as compared with 61% for average (p=0.01) and 47% for lower (p=0.00) for c; and for d, 37% for higher compared with 16% for average (p=0.01) and 19% for lower (p=0.00). This indicates students in the higher reasoning group have similar problem solving abilities between problem types, as do the lower reasoning group, just at a different performance level. However, the performance of those in the average reasoning group shows a variable difference, particularly on the upper level problem categories.

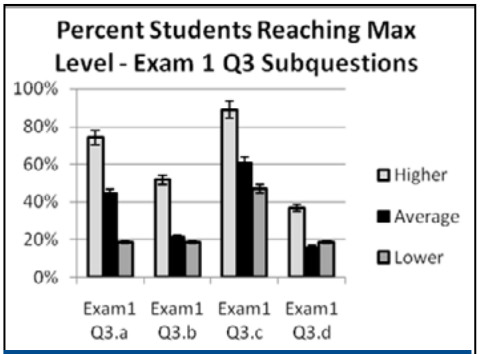


Figure 7. Comparison of subquestion scores on different aspects of projectile motion. Over 50% of higher reasoners scored at the maximum level of 2b on subquestions a and b, as well as maximum level of 1a on subquestion c. The maximum level for subquestion d was 2b.

Newton's second law, exam two, question one. The same pattern was observed for projectile motion between groups and is evident when comparing the performance for Newton's laws as shown in Fig. 8. The higher reasoning group shows a similar performance, but the results are closer than in the previous problem. However, the pattern is more dramatic given all three subquestions share the same maximum value, 2b. A gap is still observed on subquestion a, between the higher (48% students reached the maximum level) and average group (38% students at the maximum level, p=0.35); and between the higher and lower group (22%, p=0.07). Although the performance of all students is low, the gap for subquestions c and d&e is significant when comparing the higher and average groups (c: 19% and 9%, respectively, p=0.01. d&e: 11% and 4%, p=0.02) and the higher and lower groups (c: 19% and 0%, respectively, p=0.00. d&e: 11% and 0%, p=0.00). When the differences within groups for each subquestion are evaluated, a decline in critical thinking performance is observed as students' progress through the problem. Since the last subquestions ('d&e') of this problem required students to analyze their responses, this may indicate students have difficulty mentally processing the information given in the question to complete the complex procedures associated with the problem's tasks.

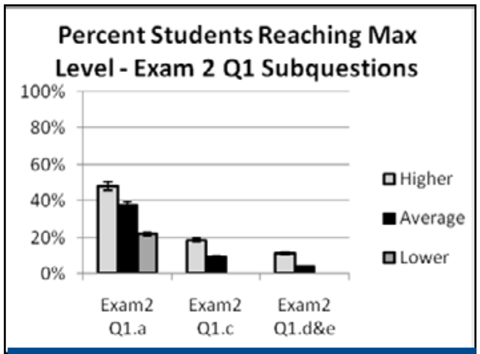


Figure 8. Comparison of subsection scores on different aspects of Newton's second law (tension and acceleration). Less than 50% of all reasoners scored at the maximum categorical level. Higher reasoners performed better than average and lower reasoners. All subquestions had a maximum level of 2b.

Newton's laws, exam two, question three. The pattern remains consistent in question three of exam two as evidenced by the comparison of performances between groups; however it is less pronounced for this problem (See Fig. 9). All three reasoning groups show similar performance when comparing data between and within groups on all three subquestions. The percentage of students reaching the maximum categorical level for subquestion b.(i,ii,iv) was low for all reasoners: 7% for higher as compared to 11% for average (p=0..70) and as compared to 6% for lower (p=0.40). The opposite is true for subquestion c, in that a high percentage of students across all reasoning levels performed at the maximum level (89% of higher reasoners compared with 84% average, p=0.89; and compared with 94% lower, p=0.14). Although the overall performance was high for subquestion e, there is still a significant gap observed between higher reasoners (96%) and average reasoners (68%, p=0.00), as well as lower reasoners (63%, p=0.00).

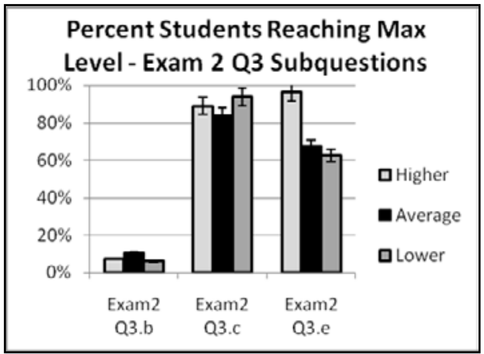


Figure 9. Comparison of subsection scores on different aspects of Newton's laws (forces, tension and acceleration). On subquestions c and e, more than 60% of all reasoners reached the maximum level, 1b. Fewer than 15% of all reasoners scored at the maximum level, 2b, for subquestion b.

Problem solving features. To conclude our results we briefly evaluated the general features of problem solving abilities for each reasoning group. A slight difference between higher reasoners (and some average reasoners) and lower reasoners was observed. Typically the lower reasoners were unable to integrate previously learned information as their solutions to the questions indicated simplistic representations of data or listed their responses in a step-by-step process. Students in the average reasoning group demonstrated higher level features, as they required fewer steps to reach a solution, using a more integrated approach. When evaluating the problem solving features of the higher reasoning group, we observed how these students demonstrated understanding of relationships between concepts as they proceeded through each subquestion. An additional feature noted, but not prominent, was some of the student's ability to properly explain concepts and their reasoning as part of their solutions.

Conclusions

The results of this study confirm that a relationship between scientific reasoning and problem solving abilities does exist. Additionally, the categorical assessment of problem solving abilities, using the Taxonomy of Physics Problems (TIPP) framework (Teodorescu et al., 2013), of students who demonstrated higher, average or lower reasoning abilities at the beginning of the introductory physics course, was successful. Since this study evaluated the first two levels of TIPP (ability to solve different types of physics problems based on retrieval of facts, performing simple calculations, integrating basic physics knowledge and those which require alternative representations through symbolization) (Teodorescu et al., 2013), it provides a good foundation in establishing a relationship between scientific reasoning and problem solving abilities.

The information shown in Figs. 4-6 regarding the individual problems provides evidence there is an observable gap in problem solving performance between the three reasoning groups. Similarly, Figs. 7-9 regarding the problem subquestions, suggests there is a range of problem solving abilities within each reasoning group.

Overall, students in the higher reasoning group outperformed students with less developed reasoning abilities on higher level physics problem solving categories and were more consistent in their performance across the three targeted problems. The performance of students with the lowest reasoning abilities also indicates consistency in their performance; however, this is at the lower level (Level 1: Retrieval) of critical thinking. Students in the average reasoning group demonstrated a higher level of problem solving abilities as compared with the lower reasoning group. This suggests students with more formal reasoning patterns may have a more advanced ability to apply physics concepts to more complex problems and engage in more advanced strategies in problem solving, which is consistent with the work of Coletta and Phillips (2010) and Teodorescu et al., (2013). In addition, these findings suggest students with lower reasoning abilities may struggle to move beyond the basic recall and execution levels in problem solving and may benefit from support mechanisms targeting scientific reasoning skills to be successful. However, more research is needed to determine if a causal relationship exists.

Understanding how students of various reasoning abilities enrolled in an introductory physics course solve different types of problems based on the levels of critical thinking will assist educators in making changes to the course design to enhance student learning. Similarly, observing where student responses fall along the problem solving continuum will assist educators in making predictions about students' problem solving abilities and success in the course. This is important because it may provide insight for educators and researchers to develop curriculum and assessments targeting a combination of reasoning and problem solving skills, as well as support mechanisms to assist students in learning how to become effective problem solvers.

Future Research

These findings expand upon an earlier pilot study comparing scientific reasoning with problem solving abilities associated with algorithmic and conceptually-based problems (Fabby & Koenig, 2014). Future work will include expanding the study to a larger sample and incorporating a wider range of physics problems that test all levels of problem solving abilities according to the TIPP Taxonomy (Teodorescu et al., 2013). As part of this expansion, a more thorough evaluation of problem-solving features will be completed, as development of future curriculum may depend upon how students learn and understand the underlying concepts.

Acknowledgements

We wish to thank the course instructors for access to their student exams. We also thank the reviewers for their useful feedback. Partial support for this work has been provided by the National Institutes of Health under Grant No. 1RC1RR028402-01.

References

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-152.

Coletta, V.P, & Phillips, J.A. (2010). Developing Thinking and Problem Solving Skills in Introductory Mechanics. *Proceedings of the Physics Education Research Conference, 1289*, 13-16.

Fabby, C., & Koenig, K. (2014). Relationship of Scientific Reasoning to Solving Different Physics Problem Types. *Proceedings from PERC Portland, OR, July 17-18, 2013.* Eds. Engelhardt, P.V., Churukian, A.D., & Jones, D.L.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence: An essay on the construction of formal operational structures.* New York: Basic Books.

Lawson, A. E. (1978). The development and validation of a classroom test of formal reasoning, *Journal of Research in Science Teaching, 15*(1), 11-24. Test used in study: Classroom Test of Scientific Reasoning, revised ed. (2000).

Marzano, R. J., & Kendall, J. S. (Eds.). (2007). *The new taxonomy of educational objectives*. Thousand Oaks: Corwin Press.

NGSS Lead States (2013). Next generation science standards: For states, by states. Achieve, Inc. on behalf of the twenty-six states and partners that collaborated on the NGSS. Retrieved from: http://www.nextgenscience.org/

Shrager, J., & Siegler, R. S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science, 9*(5), 405-410.

Teodorescu, R. E., Bennhold, C., Feldman, G., & Medsker, L. (2013). New approach to analyzing physics problems: A Taxonomy of Introductory Physics Problems. *Physical Review Special Topics-Physics Education Research, 9*(1), 010103.

Tuminaro, J. & Redish, E.F. (2007). Elements of a Cognitive Model of Physics Problem Solving: Epistemic Games, *Phys. Rev. STPER, 3*, 020101.

Carol Fabby is a doctoral candidate in Physics Education Research as part of the Educational Studies program in the College of Education at the University of Cincinnati. She received her BS and MS in Physics at the University of Cincinnati and was a three-year recipient of a Hazel F. Gabbard Doctoral Research Associate Scholarship. Her current research is investigating statistical reasoning in the introductory physics course. She is an adjunct instructor in the Department of Physics at the University of Cincinnati and is presently working on an NSF grant which focuses on developing scientific reasoning for STEM majors.

Kathleen Koenig received her PhD in Physics Education in 2004 after serving as a high school physics teacher for 6 years. She began her career in higher education at Wright State University, and she is currently an Associate Professor of STEM Education with a joint appointment in Physics at the University of Cincinnati. Her research interests include developing curriculum to improve student acquisition of scientific reasoning abilities as well as improving student success and retention in introductory STEM courses.