

## Scientific reasoning abilities of nonscience majors in physics-based courses

J. Christopher Moore\* and Louis J. Rubbo†

Department of Chemistry and Physics, Coastal Carolina University, Conway, South Carolina, 29528, USA  
(Received 12 October 2011; published 15 February 2012)

We have found that non-STEM (science, technology, engineering, and mathematics) majors taking either a conceptual physics or astronomy course at two regional comprehensive institutions score significantly lower preinstruction on the Lawson's Classroom Test of Scientific Reasoning (LCTSR) in comparison to national average STEM majors. Based on LCTSR score, the majority of non-STEM students can be classified as either concrete operational or transitional reasoners in Piaget's theory of cognitive development, whereas in the STEM population formal operational reasoners are far more prevalent. In particular, non-STEM students demonstrate significant difficulty with proportional and hypothetico-deductive reasoning. Prescores on the LCTSR are correlated with normalized learning gains on various concept inventories. The correlation is strongest for content that can be categorized as mostly theoretical, meaning a lack of directly observable exemplars, and weakest for content categorized as mostly descriptive, where directly observable exemplars are abundant. Although the implementation of research-verified, interactive engagement pedagogy can lead to gains in content knowledge, significant gains in theoretical content (such as force and energy) are more difficult with non-STEM students. We also observe no significant gains on the LCTSR without explicit instruction in scientific reasoning patterns. These results further demonstrate that differences in student populations are important when comparing normalized gains on concept inventories, and the achievement of significant gains in scientific reasoning requires a reevaluation of the traditional approach to physics for non-STEM students.

DOI: [10.1103/PhysRevSTPER.8.010106](https://doi.org/10.1103/PhysRevSTPER.8.010106)

PACS numbers: 01.40.G-

### I. INTRODUCTION

University courses in conceptual physics and astronomy typically serve as students' terminal science experience. Significant work has gone into developing research-verified pedagogical methods for preservice teachers and the algebra- and calculus-based physics courses typically populated by natural and physical science majors [1]; however, there is significantly less volume in the literature concerning the nonscience, general education population [2]. This is quickly changing, and large, repeatable gains on concept tests are being reported, specifically within the astronomy education community [3].

Student scientific reasoning and metacognition are increasingly being investigated within the physics education community, as well, though most studies use data sets consisting mostly of science majors [4,5]. Bao *et al.* demonstrate that although U.S. college-level science students perform poorly with respect to physics content knowledge relative to their Chinese peers, there is no significant difference between the two groups with respect to scientific reasoning, and that both groups demonstrate reasonable

preparation [5]. It is encouraging that students self-selecting a science major demonstrate competence in scientific reasoning at the postsecondary level. However, there is little data on the general education student population, which will comprise a greater percentage of the professional population.

Since most students enrolled in conceptual physics or astronomy will never take another formal science course, our student learning objectives should incorporate broader reasoning skills. Scientific reasoning and metacognitive development are often required for effective decision making and problem solving far outside the typical scientific context [6–8]. Furthermore, it has been shown that gains in physics content knowledge are strongly correlated to scientific reasoning [4]. In particular, reasoning and metacognition development are essential for problem solving, understanding and applying abstract concepts, and shifting between multiple representations [2,9,10]. However, non-STEM (science, technology, engineering, and mathematics) majors may enter the classroom with a disadvantage not necessarily shared by their self-selecting science major peers. Acknowledgement of the potential dramatic difference in reasoning ability is important for development of good pedagogy. Furthermore, gains in content knowledge achieved via research-verified, active-engagement curriculum may not necessarily lead to gains in scientific reasoning [11]. In fact, the content-specific education literature in other disciplines suggests that explicit intervention is necessary to improve reasoning

\*moorejc@coastal.edu

†lrubbo@coastal.edu

[2,9,12,13]. It is this explicit intervention that is currently lacking in many pedagogical models that address this student population in physics and astronomy, specifically those models practical for implementation with a large student-to-faculty ratio.

In this paper, we evaluate the postsecondary scientific reasoning abilities of non-STEM students using Lawson's Classroom Test of Scientific Reasoning (LCTSR) [14,15]. The purpose is to determine where the general education population is with respect to scientific reasoning in order to inform future pedagogies aimed at improving scientific reasoning within this group. Based on LCTSR scores, we classify students into one of three formal reasoning levels as described by Piaget's theory of cognitive development: concrete operational, transitional, and formal operational [16]. Furthermore, we investigate the types of content with which non-STEM students struggle and correlate learning gains for this content to preparation in scientific reasoning. Finally, we look at the effectiveness of "reformed" pedagogy for learning gains in scientific reasoning and discuss possible implications for instruction.

## II. BACKGROUND

In this section, we review the Piagetian levels of formal reasoning, the assessment of scientific reasoning, and the connection between reasoning level and potential gains in specific types of content knowledge. We also discuss our specific student population and the types of courses in which these students enroll.

What exactly constitutes scientific reasoning is both complex and debatable. Lawson suggests that scientific reasoning has a structure that is chiefly hypothetico-deductive in nature and consists of interrelated aspects, such as proportional reasoning, control of variables, probability reasoning, and correlation reasoning [17,18]. Inductive and deductive process are involved, with some researchers intimately linking reasoning with the process of drawing inferences from initial premises [8,19].

More recently, Kuhn has suggested that scientific reasoning is more than inductive inference, but a truth-seeking social process that involves the coordination of theory and evidence [20]. Kuhn and others specifically suggest that reasoning the process cannot be separated from prior knowledge [15,20–22]. Similarly, the learning of content and reasoning development have been linked in physics education literature [4,23]. As we will discuss further, content gains are significantly more difficult to achieve with underprepared students versus well-prepared students, and gains in reasoning only materialize with explicit intervention.

### A. Scientific reasoning and concept construction

Piaget's theory of cognitive development includes classification into two formal reasoning levels (concrete

operational and formal operational) with a transitional stage between the two [16,24,25]. Students classified as mostly concrete operational reasoners are characterized by their appropriate use of logic; however, they struggle with solving problems outside of a concrete context, demonstrating significant difficulty with abstract concepts and hypothetical tasks. Formal operational reasoners begin to think abstractly, reason logically, and draw conclusions from available information. Furthermore, unlike the concrete operational reasoner, they are able to apply appropriate logic to hypothetical situations in most contexts. In this way, formal operational reasoners can begin to think like a scientist, and specifically develop strong hypothetico-deductive reasoning. Transitional reasoners fall between the other two classifications where they find success with hypothetical tasks in some contexts. Lawson describes these levels as level 0, low level 1, and high level 1, respectively [15,26]. Lawson further describes a postformal level of reasoning, which is beyond the scope of this study. In this paper, we will use the traditional Piagetian labels.

It has been shown that the LCTSR can be used as an assessment of formal reasoning level, and its validity has been established [13–15,17,18,26]. Specifically, Lawson investigated the development of scientific reasoning and formal reasoning level in introductory college biology [15,26]. For physics, Ates *et al.* correlate formal reasoning level to conceptual understanding and problem-solving skills in introductory mechanics [23]. In both of these studies, student populations consist primarily of science majors or science education prospective teachers. In the present study, we assess formal reasoning level of non-science, general education students at the college level.

We have used the 2000 revised, multiple-choice edition of the LCTSR, which assesses reasoning patterns such as proportional reasoning, control of variables, probability reasoning, correlation reasoning, and hypothetico-deductive reasoning [15]. The LCTSR consists of 12 scenarios followed by two questions each assessing 6 different scientific reasoning patterns. Each reasoning pattern is addressed by two question pairs. One question in a pair elicits a response requiring effective use of the pattern, while the second question has the student describe the reasoning behind the response. Overall there are 12 questions designed to assess application of a reasoning pattern and 12 questions designed to evaluate the student's personal approach to that application. Requiring students to commit to a specific reason for their answers prevents them from achieving correct answers for the incorrect reasons. In order for a student to receive a "correct" score, they must answer both questions within a scenario correctly.

With respect to specific content, Lawson has shown that content can be classified into three categories: descriptive, hypothetical, and theoretical [26]. Content categorized as

descriptive includes concepts having directly observable exemplars. Examples include more concrete content requiring no more than rote memorization and/or directly observable content such as the conservation of mass, bulb brightness in circuits, reflection and refraction in optics, and moon phases. Content categorized as theoretical includes concepts without directly observable exemplars. Examples within physics include the more abstract concepts of force, energy, and vector fields. Hypothetical content typically involves observable exemplars, however, typically over a time or space scale that makes direct observation impossible. Lawson demonstrates that concrete operational reasoners can learn descriptive content with relative ease, whereas they struggle with theoretical and hypothetical content. Formal operational reasoners do relatively well with most content.

### B. Student population and course structure

The data set used in this study consists of students enrolled in either a conceptual physics or astronomy course with one of the authors during the past three years. The conceptual physics course is based on *Physics by Inquiry* (PbI), which is a guided-inquiry approach to content “in which the primary emphasis is on discovering rather than memorizing and in which teaching is by questioning rather than by telling” [27]. Over the past three years, this course was taught using a large-enrollment implementation similar to that reported by Scherr [28]. For two years, this course was taught at Longwood University in Farmville, VA, which is a primarily undergraduate, regional comprehensive institution. For one year this course was taught at Coastal Carolina University (CCU) in Conway, SC. Coastal Carolina is a similar comprehensive, regional institution. At Longwood, this course was taught in a traditional lecture room with 50–70 students and one instructor. Like Scherr’s population, students proceeded through PbI materials in groups, but with whole-class “checkouts” rather than instructor-intensive, individual group checkouts. At CCU, the course was taught in a lecture room designed for the Student Centered Activities for Large Enrollment University Physics (SCALE-UP) model [29]. A SCALE-UP course incorporates the high-impact practice of collaborative assignments and projects by fusing lecture, laboratory, and recitation into a single entity. Although the courses at Longwood were taught in a traditional classroom setting, the principles of the SCALE-UP model were implemented as much as was feasible.

The conceptual astronomy course was taught over the past three years at CCU in the same SCALE-UP classroom used for the conceptual physics course. An implementation of *Lecture Tutorials in Introductory Astronomy* and Peer Instruction was used with this course [30–32]. Typical classes had enrollments of 30–40 students with two instructors.

### III. SCIENTIFIC REASONING IN THE NON-STEM POPULATION

We have found that students in our conceptual physics and astronomy courses score significantly lower on the LCTSR compared to students enrolled in courses typically populated with science majors. Table I shows average LCTSR preinstruction scores ( $N = 1208$ , average = 75%) for science and engineering majors enrolled in a calculus-based introductory physics course, as reported by Bao, *et al.* [5]. The LCTSR was also administered to students taking a conceptual physics or astronomy course with one of the authors during the past three years. As shown in Table I, this population of students scores significantly lower ( $N = 109$ , average = 54%). We found no significant difference between students in the conceptual physics and astronomy courses, or between students at the two institutions.

That students in STEM majors demonstrate stronger scientific reasoning ability is not surprising, since most students typically choose their major based on their strengths. However, such a dramatic difference in reasoning ability between STEM and non-STEM students may contribute to disparities in effectiveness of reformed physics pedagogies. What works in calculus-based physics courses with natural and physical science students may not work in the general education, conceptual physics course.

To better understand our student population, we have investigated non-STEM student weaknesses with respect to scientific reasoning. Specifically, we looked at the formal reasoning level for the average non-STEM student and determined which scientific reasoning patterns presented the most difficulty for this population.

#### A. Formal reasoning levels

Using individual student scores on the LCTSR, we classified students into three formal reasoning categories: concrete operational (CO), transitional (T), and formal operational (FO). As described by Lawson, students scoring below 25% on the LCTSR were classified as concrete operational reasoners, students scoring between 25% and 58% were classified as transitional reasoners, and students scoring above 58% were classified as formal operational reasoners [15]. As previously mentioned, a student must correctly answer both questions within a scenario in order to receive credit. There are 24 questions on the LCTSR, with 12 different scenarios assessing 6 different reasoning

TABLE I. Average LCTSR scores for STEM and non-STEM majors.

	LCTSR %	$N$	Standard deviation
STEM	75	1208	18
Non-STEM	54	109	17

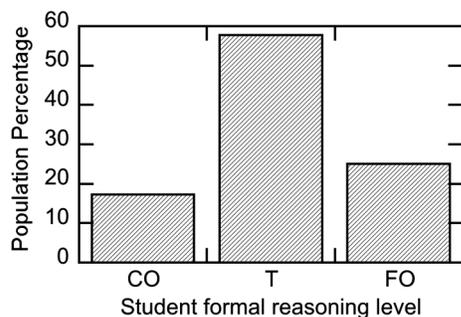


FIG. 1. Distribution of formal reasoning level for non-STEM students enrolled in conceptual physics or astronomy. Formal reasoning level was determined by the LCTSR. Concrete operational (CO), transitional (T), and formal operational (FO) reasoning levels are shown.

patterns. Percentages were calculated based on the 12 scenarios.

Figure 1 shows the distribution of non-STEM students within Piagetian formal reasoning levels. A significant majority of non-STEM students (56%) are classified as transitional reasoners. This observation is consistent with previous studies of the general education population in introductory biology courses for the nonmajor [33]. It should be noted that the majority of science majors in an introductory, calculus- or algebra-based physics course would, on average, be classified within the formal operational category, which is also observed in the biology education literature for the STEM-based biology introductory courses [5,34].

### B. Analysis of specific reasoning patterns

Figure 2 shows the population averages for specific scientific reasoning patterns as assessed by the LCTSR. Students within the observed population demonstrated

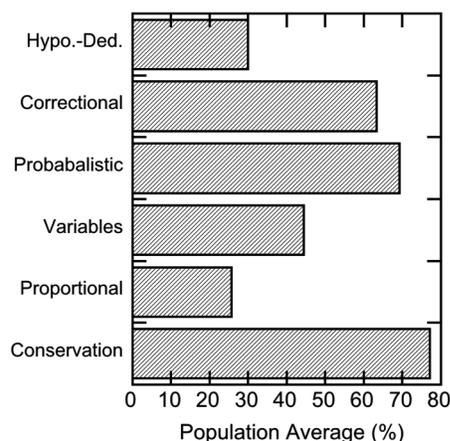


FIG. 2. Population averages on specific scientific reasoning patterns as assessed by the LCTSR. Significant difficulty is observed with proportional reasoning, isolation of variables, and hypothetico-deductive reasoning.

significant difficulty with proportional reasoning, isolation of variables, and hypothetico-deductive reasoning. Of particular interest, scores on LCTSR questions designed to test application of hypothetico-deductive reasoning, which can arguably be called the “scientific method,” average to an abysmal 30%. Students within this population demonstrated the poorest performance on proportional reasoning (25%), which could be attributable to poor preparation in mathematics. Surprisingly, most students demonstrate some proficiency with correctional and probabilistic reasoning. The dramatic difference between performance on probabilistic or correctional reasoning and proportional reasoning is puzzling. It is possible that the standard secondary school curriculum includes more explicit instruction on these reasoning patterns.

As instructors we must be careful not to assume preexisting knowledge, specifically with respect to content requiring use of these patterns. For example, choice of appropriate scale is essential when graphing data and/or moving between multiple representations. Furthermore, qualitative approaches to physical systems may be more difficult for this population, due to a lack of proficiency with proportion. It is evident that more explicit instruction in scientific reasoning is necessary, specifically with respect to hypothesis construction and testing, proportions, and isolation of variables.

## IV. IMPLICATIONS FOR INSTRUCTION

Analysis of students’ scientific reasoning level introduces several implications for instruction. In particular, we expect transitional and concrete operational reasoners to have difficulty with theoretical content. This difficulty could result in low learning gains on concept inventories for these students when compared to formal operational reasoners receiving similar instruction. As others have previously demonstrated, preexisting knowledge and skill sets do contribute to the maximum achievable learning gain in physics [4].

In this section, we look at the correlation between preexisting scientific reasoning ability and learning gains for various classifications of content. Specifically, is there a stronger correlation between LCTSR prescore and normalized learning gain for more theoretical content in comparison to descriptive content? We also discuss whether content-driven pedagogies significantly contribute to gains in scientific reasoning. Is a guided-inquiry-based approach to learning science sufficient for the development of scientific reasoning?

### A. Correlation between reasoning and content knowledge gains

Lawson demonstrated a connection between concept construction and developmental level in college biology [26]. We expect a similar connection to be prevalent in a physics context. Specifically, we expect to see a correlation

between scientific reasoning and knowledge gains in a physics-based context.

Coletta and Phillips observed a strong correlation between normalized gain on the Force Concept Inventory (FCI) and preinstruction LCTSR scores [4]. During assessment for our conceptual physics courses over the past two years, we have observed similar strong correlations between preinstruction LCTSR scores and normalized gain on two concept inventories, the Determining and Interpreting Resistive Electric Circuits Concept Test (DIRECT) and the Test for Understanding Graphs-Kinematics (TUG-K) [35,36]. As shown in Fig. 3, a strong correlation is seen for content requiring higher-order and more abstract reasoning. With a slope of linear fit of 0.64 and  $r = 0.59$ , the correlation between TUG-K normalized gain and LCTSR score is similar to that seen for the FCI and stronger than the correlation observed for the DIRECT assessment [slope = 0.45 and  $r = 0.50$  (see Fig. 4)].

The TUG-K tests a student's ability to move between multiple representations, which rely on higher-order and more abstract thinking. The FCI assesses a student's knowledge and application of the abstract concept of force. These concepts can be classified as mostly theoretical, requiring advanced reasoning development to achieve success. A weaker correlation between DIRECT gains and LCTSR scores could be because strong scores are possible on DIRECT via good observation and retention from well-

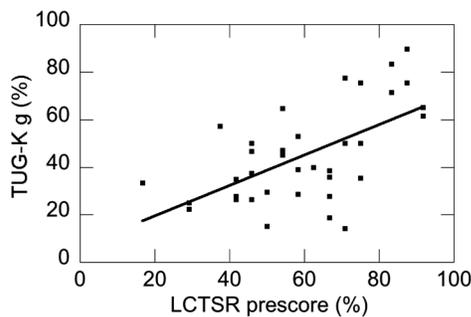


FIG. 3. Normalized learning gains on TUG-K versus LCTSR prescore. The solid line is the line of best fit (slope = 0.64 and  $r = 0.59$ ).

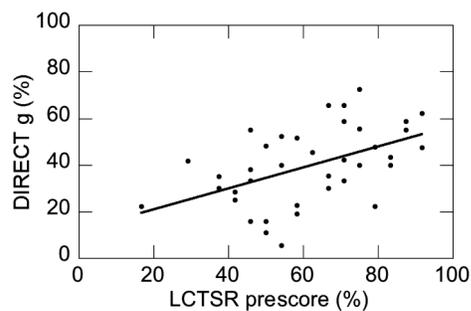


FIG. 4. Normalized learning gains on DIRECT versus LCTSR prescore. The solid line is the line of best fit (slope = 0.45 and  $r = 0.50$ ).

designed inquiry-based activities; many of the questions are based on content with explicit, concrete exemplars that are directly observed during the course. This suggests that if we wish to push our nonscience students past the lower three levels of Bloom's taxonomy of educational objectives, then we may need our courses to focus explicitly on scientific reasoning early and often [37].

It should be noted that we use a specific pedagogical approach within our courses. Therefore, postinstruction gains themselves may or may not be indicative of what would be found in other courses following different pedagogies. However, the correlations between LCTSR prescore and normalized gain on the concept inventories discussed are still relevant within the context of a qualitative discussion on the general link between scientific reasoning ability and content knowledge "learning potential" within specific categories of content.

## B. Reformed pedagogy and gains in reasoning

Considering the link between scientific reasoning and learning gains observed within our population, and the population's relatively low preinstruction LCTSR scores, we might expect difficulty achieving meaningful gains. However, substantial gains in content knowledge can still be obtained in conceptual physics and astronomy courses, especially when those courses are designed around a research-verified, active-engagement curriculum. Table II shows average normalized learning gains on DIRECT, TUG-K, and the Star Properties Concept Inventory (SPCI) [38] for a subset of students enrolled in our courses over the past three years. Although lower than reported for students completing some active-engagement algebra- and calculus-based courses, these gains are still significant. Even though we have been relatively successful with content, we have failed to improve reasoning ability. Average normalized gains on the LCTSR for both physics and astronomy students are essentially equivalent to zero.

The lack of significant gains in scientific reasoning is particularly surprising for the conceptual physics course, which via PbI is completely designed around the process of scientific inquiry. However, it should be pointed out that we have implemented an adaptation of PbI that strays in some significant ways from the intentions of the curriculum

TABLE II. Average normalized gains on three conceptual inventories and the LCTSR. Students taking the conceptual physics course were assessed via DIRECT and TUG-K. Students taking conceptual astronomy were assessed via the SPCI. Matched data from the LCTSR were obtained from both courses.

	$g$	$N$	Standard deviation
DIRECT	0.38	40	0.04
TUG-K	0.42	38	0.05
SPCI	0.39	36	0.05
LCTSR	0.06	62	0.08

designers. Therefore, the data should not be interpreted as condemnation of any particular pedagogy. Furthermore, this does not suggest that gains in reasoning are unachievable. The content-specific education literature in other disciplines suggests that explicit intervention is necessary to improve reasoning [2,9,12,13]. Once again, our main point is certainly not that some specific pedagogy or collection of pedagogies fails to impact reasoning, but that any classroom approach could without explicit intervention.

## V. INFORMING FUTURE PEDAGOGIES

There are three main conclusions from this work that have implications for physics and astronomy instruction with this population: (1) reformed pedagogy focused on content alone is not necessarily sufficient to achieve gains in scientific reasoning; (2) scientific reasoning can be strongly correlated to gains in content knowledge, especially content categorized as theoretical; and (3) students entering our conceptual physics and astronomy courses demonstrate poor preparation in some scientific reasoning patterns. A dramatic difference in reasoning ability between STEM and non-STEM students may contribute to disparities in effectiveness of reformed physics pedagogies. What works in calculus-based physics courses with natural and physical science students may not work in the general education, conceptual physics course.

In this section we will briefly discuss how these findings can be used to inform the development of new pedagogies. Specifically, we discuss the necessity of making scientific reasoning explicit within our conceptual physics and astronomy courses and the affect topic sequence could have on student learning.

### A. Making scientific reasoning explicit

The content-specific education literature in other disciplines suggests that explicit intervention is necessary to improve reasoning [2,9,12,13]. In fact, we are beginning to see significantly larger gains in scientific reasoning via explicit instruction during our most recent courses, though these observations are preliminary.

For example, Lawson presents a series of activities that lead students through the process of constructing good “if ... and ... then” (IAT) statements in various fields of knowledge, though with a focus on applications within the biology content of his courses [13,15,39]. In this way, hypothetico-deductive reasoning is being made explicit. We have begun introducing these types of activities within our course with preliminary success. As an example, PbI has students design an experiment that would address the question “does a light bulb use up current?” We formalize the process by forcing students to construct an appropriate IAT statement that is specifically designed to falsify a claim. Based on observations from a test class consisting of 15 students from both the physics major and general edu-

TABLE III. Average normalized gains on two conceptual inventories and the LCTSR for students participating in a conceptual physics course with and without explicit reasoning intervention.

	$g$	$N$	Standard deviation
DIRECT	0.41	15	0.04
TUG-K	0.55	14	0.05
LCTSR	0.68	14	0.08

cation population, we have found that students initially struggle with IAT statement construction but gradually improve. We combined these activities with activities explicitly targeting other reasoning patterns (proportional, control of variables, probability reasoning, correlation reasoning) all mixed within the actual content as laid out in PbI.

This explicit approach to reasoning development appears to result in increased gains on the LCTSR and TUG-K. As shown in Table III, an average normalized gain of 68% ( $n = 14$ ) was achieved on the LCTSR in comparison to 11% ( $n = 42$ ) for previous courses taught similarly with respect to content, though lacking explicit reasoning intervention. Furthermore, content gains as measured by TUG-K were significantly higher for the explicit instruction test group, also shown in Table III. Interestingly, there is no significant difference in gains on the DIRECT assessment between the nonexplicit and explicit groups. This is consistent with our observation of a weaker correlation between LCTSR prescore and learning gains on the DIRECT instrument. Although our preliminary sample is small, these preliminary results are encouraging and, combined with results from the education literature in other fields, point to a potential high-reward approach. We are also actively developing activities targeting reasoning patterns for the conceptual astronomy course.

### B. Could topic sequence affect student learning gains?

Topic sequence could have an effect on student learning gains, even with explicit instruction in scientific reasoning patterns, especially with the conceptual physics course. An examination of several conceptual physics and physical science textbooks shows that the average university course begins with topics in mechanics, specifically motion, force, and energy. These more theoretical concepts are the hardest in which to achieve significant learning gains with this population of students; however, they are typically the first topics covered in the traditional conceptual physics course.

A topic sequence where more concrete content is initially introduced along with explicit instruction in scientific reasoning may lead to better performance in theoretical content later in the course. As an example, much of the content within the topics of circuits and optics has directly observable exemplars, which makes them ideal candidates for introductory material. With regard to circuits, McDermott has shown that no formal instruction in

electrostatics or field theories is needed to successfully build a predictive model [40]. It may be beneficial for students in conceptual astronomy to begin with observation-based activities and physical model building, such as the celestial sphere. Gradually increasing the amount of hypothetical and theoretical content as reasoning develops could build student confidence in handling physics while better preparing them for theoretical content.

## VI. SUMMARY

In summary, we have found that non-STEM majors taking either a conceptual physics or astronomy course at two regional comprehensive institutions score significantly lower preinstruction on the LCTSR in comparison to STEM majors. Furthermore, there is a strong correlation between preinstruction LCTSR scores and normalized gains on concept inventories. The correlation is strongest for content that can be categorized as mostly theoretical, meaning a lack of directly observable exemplars, and weakest for content categorized as mostly descriptive, where directly observable exemplars are abundant.

Although the implementation of research-verified, interactive engagement pedagogy can lead to gains in content

knowledge, significant gains in theoretical content (such as force and energy) are more difficult with this population of students. We also observe no significant gains on the LCTSR without explicit instruction in scientific reasoning patterns. This has several implications for instruction, such as the necessity for explicit instruction and the potential need for a reassessment of the canonical sequence of topics in conceptual physics and astronomy.

These results further demonstrate that differences in student populations are important when comparing normalized gains on concept inventories, and the achievement of significant gains in scientific reasoning requires a reevaluation of the traditional approach to physics for non-STEM populations.

## ACKNOWLEDGMENTS

The authors would like to thank Michelle Parry, Associate Professor of Physics at Longwood University, for helpful discussions about the development of the conceptual physics course and scientific reasoning. We would also like to thank Janelle Bailey, Assistant Professor of Science Education at the University of Nevada, Las Vegas, for providing initial analysis of our SPCI raw data.

- 
- [1] L.C. McDermott and E.F. Redish, Resource letter: PER-1: Physics education research, *Am. J. Phys.* **67**, 755 (1999).
  - [2] E. Etkina and J.P. Mestre, *Implications of Learning Research for Teaching Science to Non-Science Majors* (SENCER, Harrisburg, PA, 2004).
  - [3] Janelle M. Bailey and Timothy F. Slater, A review of astronomy education research, *Astron. Educ. Rev.* **2**, 20 (2003).
  - [4] Vincent P. Coletta and Jeffrey A. Phillips, Interpreting FCI scores: Normalized gain, preinstruction scores, and scientific reasoning ability, *Am. J. Phys.* **73**, 1172 (2005).
  - [5] Lei Bao *et al.*, Learning and scientific reasoning, *Science* **323**, 586 (2009).
  - [6] F. Reif and J.H. Larkin, Cognition in scientific and everyday domains: Comparisons and learning implications, *J. Res. Sci. Teach.* **28**, 733 (1991).
  - [7] A.F. Greenhoot, G. Semb, J. Colombo, and T. Schreiber, Prior beliefs and methodological concepts in scientific reasoning, *Appl. Cogn. Psychol.* **18**, 203 (2004).
  - [8] *Reasoning, Necessity, and Logic: Developmental Perspectives*, edited by W.F. Overton (Lawrence Erlbaum Assoc., Hillsdale, NJ, 1990).
  - [9] B.Y. White and J.R. Frederiksen, Inquiry, modeling, and metacognition: Making science accessible to all students, *Cogn. Instr.* **16**, 3 (1998).
  - [10] P. Kohl and N.D. Finkelstein, Patterns of multiple representation use by experts and novices during physics problem solving, *Phys. Rev. ST Phys. Educ. Res.* **4**, 010111 (2008).
  - [11] N.R. Pearsall, J.E.J. Skipper, and J.J. Mintzes, Knowledge restructuring in the life sciences: A longitudinal study of conceptual change in biology, *Sci. Educ.* **81**, 193 (1997).
  - [12] L.M. Blank, A metacognitive learning cycle: A better warranty for student understanding?, *Sci. Educ.* **84**, 486 (2000).
  - [13] Anton E. Lawson, The generality of hypothetico-deductive reasoning: Making scientific thinking explicit, *Am. Biol. Teach.* **62**, 482 (2000).
  - [14] A.E. Lawson, The development and validation of a classroom test of formal reasoning, *J. Res. Sci. Teach.* **15**, 11 (1978).
  - [15] A.E. Lawson, B. Clark, E. Cramer-Meldrum, K.A. Falconer, J.M. Sequist, and Y. Kwon, Development of scientific reasoning in college biology: Do two levels of general hypothesis-testing skills exist?, *J. Res. Sci. Teach.* **37**, 81 (2000).
  - [16] B. Inhelder and J. Piaget, *The Growth of Logical Thinking from Childhood to Adolescence* (Basic, New York, 1958).
  - [17] A.E. Lawson, The nature of advanced reasoning and science instruction, *J. Res. Sci. Teach.* **19**, 743 (1982).
  - [18] A.E. Lawson, What is the role of induction and deduction in reasoning and scientific inquiry?, *J. Res. Sci. Teach.* **42**, 716 (2005).

- [19] K.J. Holyoak and R.G. Morrison, *Thinking and Reasoning: A Reader's Guide* (Cambridge University Press, Cambridge, England, 2005).
- [20] D. Kuhn, in *Blackwell Handbook of Childhood Cognitive Development*, edited by U. Goswami (Wiley-Blackwell, Malden, MA, 2004), pp. 371–393.
- [21] L. Schauble, The development of scientific reasoning in knowledge-rich contexts, *Dev. Psychol.* **32**, 102 (1996).
- [22] K. Dunbar, How do scientists think in the real world: Implications for science education, *J. Appl. Dev. Psychol.* **21**, 49 (2000).
- [23] S. Ates and E. Cataloglu, The effects of students' reasoning abilities on conceptual understanding and problem-solving skills in introductory mechanics, *Eur. J. Phys.* **28**, 1161 (2007).
- [24] H. Ginsburg and S. Opper, *Piaget's Theory of Intellectual Development* (Prentice-Hall, Upper Saddle River, NJ, 1979).
- [25] J.W. Santrock, *A Topical Approach to Life Span Development* (McGraw-Hill, New York, 2008).
- [26] A.E. Lawson, S. Alkhoury, R. Benford, B.R. Clark, and K.A. Falconer, What kinds of scientific concepts exist? Concept construction and intellectual development in college biology, *J. Res. Sci. Teach.* **37**, 996 (2000).
- [27] L.C. McDermott *et al.*, *Physics by Inquiry* (John Wiley and Sons, New York, 1996), Vol. I.
- [28] R.E. Scherr, An implementation of Physics by Inquiry in a large-enrollment class, *Phys. Teach.* **41**, 113 (2003).
- [29] R. Beichner *et al.*, Case study of the physics component of an integrated curriculum, *Am. J. Phys.* **67**, S16 (1999).
- [30] E.E. Prather, T.F. Slater, J.P. Adams, J.M. Bailey, L.V. Jones, and J.A. Dostal, Research on a lecture-tutorial approach to teaching introductory astronomy for non-Science majors, *Astron. Educ. Rev.* **3**, 122 (2004).
- [31] C.H. Crouch and E. Mazur, Peer instruction: Ten years of experience and results, *Am. J. Phys.* **69**, 970 (2001).
- [32] E. Mazur, Farewell, Lecture?, *Science* **323**, 50 (2009).
- [33] M.A. Johnson and A.E. Lawson, What are the relative effects of reasoning ability and prior knowledge on biology achievement in expository and inquiry classes?, *J. Res. Sci. Teach.* **35**, 89 (1998).
- [34] M.S. Schen, Ph.D. thesis, The Ohio State University, 2007.
- [35] Paula Vetter Engelharda and Robert J. Beichner, Students' understanding of direct current resistive electrical circuits, *Am. J. Phys.* **72**, 98 (2004).
- [36] Robert J. Beichner, Testing student interpretation of kinematics graphs, *Am. J. Phys.* **62**, 750 (1994).
- [37] *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*, edited by L.W. Anderson *et al.* (Allyn & Bacon, Boston, MA, 2001).
- [38] Janelle M. Bailey, Development of a concept inventory to assess students' understanding and reasoning difficulties about the properties and formation of stars, *Astron. Educ. Rev.* **6**, 133 (2007).
- [39] A.E. Lawson, A scientific approach to teaching about evolution and special creation, *Am. Biol. Teach.* **61**, 266 (1999).
- [40] L.C. McDermott, P.S. Shaffer, and C.P. Constantinou, Preparing teachers to teach physics and physical science by inquiry, *Phys. Educ.* **35**, 411 (2000).