Development and evaluation of introductory physics lab curriculum to promote scientific reasoning abilities

Kathleen KOENIG¹, Krista E WOOD² and Lei BAO³

¹Physics Department, University of Cincinnati, Cincinnati, OH 45221 USA
²Math, Physics and Computer Science Department, University of Cincinnati-Blue Ash, Blue Ash, OH 45236 USA
³Physics Department, The Ohio State University, Columbus, OH 43210 USA

Abstract. Research indicates students with more formal reasoning patterns are more proficient learners, while certain critical thinking abilities are important for living in modern society. Unfortunately, college courses do not typically promote these skills. In order to address this concern, we have developed a first semester introductory physics lab curriculum in which students engage in experimental design with emphasis placed on the use of evidence-based reasoning in making decisions in the lab and in report writing. This paper describes the underlying framework on which the curriculum is based and presents data which demonstrates significant shifts in student development of certain aspects of scientific reasoning.

1. Introduction

Addressing complex societal problems requires STEM workers with expertise such that they are able to recognize, understand, and effectively reason out solutions using principles of their disciplines. As a result, education goals have shifted from content drilling towards fostering higher end skills including reasoning, creativity, and open problem solving [1]. Consensus for defining these skills converges on 3 broad clusters: cognitive, interpersonal, and intrapersonal skills that include a rich set of sub-dimensions.

Within the cognitive domain, multiple competencies have been proposed, including non-routine problem solving, systems thinking, critical thinking, computational and information literacy, reasoning and argumentation, and innovation [2].

There are many examples of business and academic organizations with similar lists of necessary 21st Century skills [3]. In a Forbes Leadership Column, Myatt [4] indicates essential elements of critical thinking that support sound decision making, including the ability to know where and how to get complete data, discern and filter input, synthesize large amounts of information, understand that not all input weighs the same, take bias into consideration, apply context and meaning to the data for one’s own purposes, and finally, refine information through analysis such that a final decision is based on knowledge that has a high degree of accuracy.

Although many descriptions exist for what qualifies as 21st Century skills, critical thinking is the most commonly noted and widely studied. Various definitions exist, with common underlying principles. Broadly defined, critical thinking applies the cognitive skills or strategies that increase the probability of a desirable outcome. It is the thinking involved in solving problems, formulating inferences, calculating likelihoods, and making decisions [5]. It is the “reasonable reflective thinking focused on deciding what to believe or do”. [6] Critical thinking is recognized as a way to understand and evaluate subject matter; producing reliable knowledge and improving thinking itself [7].
A significant part of critical thinking depends on the ability to consciously control the coordination of claims and evidence [8]. This process of theory–evidence coordination (TEC) represents an integrative framework of reasoning where new knowledge is constructed through the intersection of student’s existing theories, data-driven outcomes, and scientifically accepted theories. This coordination for building new knowledge involves a process for considering these types of evidence as distinct, while making thoughtful connections between them. It also involves the ability to consider the potential impact of unknown, but possible causal factors. These competencies are crucial as they represent the kinds of reasoning necessary for understanding the physical world as they allow for predictions, inferences, and explanations [9,10]. It is the thinking needed by all citizens, whether or not in STEM careers, for making inferences about causality, which is essential for informed decision making.

2. Rationale for promoting theory-evidence coordination abilities in college students

Research has shown that the areas of reasoning that support theory–evidence coordination (TEC), such as abilities in controlling variables, data analytics, and causal decision-making, are lacking in students [11,12]. Other more broad areas of reasoning related to TEC have also been found lacking, including the ability to reason effectively in both science-related and everyday life contexts [13,14], to evaluate evidence independent of prior knowledge [14], and to consider alternative hypotheses [15]. For example, Lawson [16] found that ~50% of intro biology students do not use formal reasoning, including the ability to develop hypotheses, control variables, and design experiments; all necessary for inquiry. Others found undergraduates have difficulty developing evidence-based arguments and differentiating between and linking evidence with claims [17,18]. Student reasoning is further weakened by an inability to distinguish between theoretical ideas and empirical evidence, with little to no reflection on prior knowledge. Students also tend to dismiss inconsistent information, such that reasoning becomes selective or self-serving as students approach tasks with the inclination to verify rather than disconfirm [19].

Although a significant amount of research has shown that students and post-college adults have not developed basic reasoning subskills believed to be essential for TEC [20], research has also shown that certain curricular interventions can significantly advance these abilities [21]. In our own research, we found that traditional courses do not significantly develop these abilities, with pre-to-post-reasoning test gains of 1%-2% using Lawson’s Classroom Test of Scientific Reasoning [22], while inquiry-based reasoning-targeted courses have been found to have gains as high as 15% [23,24]. It is this research that prompted us to develop a one semester introductory physics lab curriculum that explicitly targets student abilities in scientific reasoning. The students targeted are enrolled in either an algebra- or calculus-based physics course. The former involves primarily health science majors and the latter involves primarily engineering majors.

3. Promoting theory-evidence coordination in the lab setting

There is an important thread of research on scientific reasoning on which our lab curriculum is grounded. It involves Kuhn’s work on multivariate causal inference and theory evidence coordination (TEC), which has broadened the field of study of scientific reasoning to go beyond investigating student abilities in controlling variables and engaging in inductive causal inferences, which had previously been the primary areas of research [8,25]. Kuhn claims that reasoning is a conscious and purposeful process for revising ideas and generating new understanding in light of evidence. This process, known as theory–evidence coordination, represents an integrative framework of reasoning that requires questioning existing theories, identifying alternative explanations, seeking and validating evidence (both supporting and contradictory), and evaluating and deciding explanations based on evidence. New knowledge is constructed through the intersection of students’ existing theories (including misconceptions), data-driven outcomes (covariation data established through controlled experiments), and scientifically accepted theories. This coordination for building new knowledge involves a process for considering these types of evidence to form a network of meaningful connections among them and between evidence and explanations. It also involves the ability to consider the potential impact of unknown, but possible
causal factors as components and relations to form new evidence and explanations. These competencies are crucial as they represent the kinds of reasoning necessary for understanding the physical world as they allow for predictions, inferences, and explanations in cyclic inquiry processes [9,10]. Kuhn’s work emphasizes the multivariable nature of causal relations embedded within a variety of reasoning contexts, and has demonstrated a lack of effective multivariable reasoning among children and post-college adults in coordinating evidence with explanations.

It is the synthesis of Kuhn’s work, combined with our previous research on college students’ development of essential scientific reasoning subskills, including control of variables, data analytics, and causal decision-making, which comprise the framework on which our lab curriculum is based (see Figure 1). The framework guides the development of our lab activities which promote the synthesis of both experimental evidence and theoretical hypotheses to develop integrative generalized understanding [26]. Within our lab curriculum, students complete 7 multi-week investigations during the semester, with each following the framework shown in Figure 1. These investigations include the following contexts: pendulum, projectile motion, Newton’s Second Law, simple harmonic motion, linear momentum, rotation, and a windmill design challenge. Each investigation involves an inquiry cycle (IBSE approach) that begins with a research question, such as What impacts the period of the pendulum?, followed by a brainstorming process that elicits students’ prior knowledge (theories). From these, students generate hypotheses, which are then tested in a lab setting to provide evidence to support or refute each. In order to generate valid experimental evidence and use the evidence for formal evaluation of hypotheses, students must have proficiency in multi-variable evidence-based reasoning, which depends on three scientific reasoning subskills including control of variables, data analytics, and causal decision-making. Using these skills, students generate valid evidence through controlled experiments that enable them to create and evaluate covariation relations from multi-variable data. From this evidence, students must identify causal relations to confirm or disconfirm hypotheses, or revise their existing theories. Students continue the cyclic process of generating and testing hypotheses until they reach satisfactory consistency between the best evaluated hypotheses and supporting evidence. When consistency cannot be satisfied within known evidence and hypotheses, students may conclude indeterminacy or open new ground in both hypothesis searching and experimental exploration.

In order to avoid inconclusive outcomes under the goal of developing scientific knowledge, students at this stage of the investigative process are guided in the curriculum to coordinate their outcomes with scientifically accepted theory. This may lead to new cycles of hypothesis testing, thereby generating new knowledge about the concept under study. This process often relies on the coordination of several distinct types of hypotheses and evidence: (a) student’s existing ideas (including bias and
misconceptions), (b) empirical evidence that may have been misinterpreted to support misconceptions and bias, (c) data patterns that don’t necessarily constitute valid evidence, (d) claims made in the lab setting supported by valid evidence, and (e) scientifically accepted theories. Because a successful TEC process requires questioning existing theories, seeking valid evidence, and eliminating inconsistent explanations, students may move either direction around the TEC framework shown in Figure 1 throughout the lab investigation [8,26].

A primary goal of the curriculum is to help students understand that multivariable equations represent a single multivariable relation construct, which may manifest itself in different forms of functionality under different controlled covariation conditions. The curriculum, therefore, engages students in a progression sequence of deliberate practice, with investigations increasing in complexity across the course, such that students come to (1) understand the range of conditions for which various relationships in multivariable expressions may exist and evolve, and (2) appropriately interpret the causal links among the relations for hypothesis evaluation under different COV settings.

In addition, we believe that in order to produce valid evidence to support the multi-variable evidence-based reasoning part of the TEC process, students need to master a rich set of scientific reasoning subskills, including: (1) control of variable (COV) reasoning, which includes the foundational skills whereby students identify, construct, and evaluate valid COV conditions that form the basis for any valid covariation data and relations in multivariable settings; (2) data analytic skills, which extend the COV reasoning to identify, manipulate, and evaluate covariation relations from multivariable data and experimental settings; and (3) causal decision-making reasoning skills, which involve relating valid evidence (covariation data and relations) with given or hypothesized theoretical claims for explanatory or predictive evaluation under different conditions, thereby serving as the conclusion stage of theory-evidence coordination. Due to the importance of this set of subskills, targeted instruction and deliberate practice using hypothetical scenarios are provided through the pre-lab activities. These exercises increase in complexity across the term and examples can be found elsewhere [26]. Also, due to the importance of these subskills, we have designed assessments to measure student development of scientific reasoning across the course, and outcomes are used to continually improve the curriculum [11,27].

4. Example demonstrating how TEC is promoted in the lab curriculum

In this section, we describe one of our lab activities in order to demonstrate how students engage in the TEC process. The guiding research question for this example activity is “What impacts the period of a pendulum?” Students are asked to brainstorm possible factors, thereby eliciting their own ideas (theories), which is one part of the TEC process shown in Figure 1. For this investigation, nearly all students expect angle of release, mass of pendulum bob, and length of pendulum to impact the period. Students are subsequently led through guided questions printed in the lab instructions to conduct controlled experiments to test each of these possible factors. Each experiment generates data which is plotted to indicate the relationship between the independent and dependent variables. Students see firsthand that the mass of the pendulum bob does not impact the period, as indicated by a period versus mass graph that is essentially a horizontal line. Similarly, students see that a period versus angle of release graph yields no dependency for small angles, such as under 20 degrees. The graph for period versus pendulum length, however, indicates a power relationship exists. Figure 2 shows a typical outcome for this experiment in our lab course. It should be noted that students when left to their own devices will only collect data for shorter lengths and therefore are likely to claim that a linear relationship exists between period and length. Instructors watch for this and intervene, when appropriate. The lab equipment is such that students can test a pendulum length of up to 1.8 m so students will be asked to test the longer lengths if their instructor notices that they haven’t done so. Once students have generated an appropriate graph, they conduct a curve fit in Excel to determine the mathematical relationship. Students are then guided through question prompts to rewrite the mathematical model (see Figure 2 as shown on the graph), in terms of the physical quantities under investigation. For the case shown in Figure 2, students would indicate the mathematical relationship between period and length as \( T = \)}
2.01√\(l\), where \(T\) is period and \(l\) is pendulum length. At this point, students are guided to connect their own theories and the evidence (i.e. experimental mathematical model) generated in the lab setting with the scientifically accepted theoretical model found in their textbook. Here, students must determine the physical meaning for the constant in their model in light of the theoretical model. In this example, students would come to recognize that the constant of 2.01 in their experimental model is essentially \(2\pi/\sqrt{g}\) when compared to the theoretical model. Through the use of additional question prompts, students are guided to discuss similarities as well as reconcile differences between the two models, thereby completing the TEC process.

**Figure 2.** Example of student data outcomes for the pendulum lab.

5. **Research on student abilities in coordinating evidence with theory**

Our research around the lab curriculum involves two areas including its impact on (1) student development of the subskills which support multi-variable evidence-based reasoning as shown in Figure 1 and (2) student abilities for engaging in the complete TEC process. In order to address the first research area, we use an assessment tool we developed for this purpose, iSTAR, *Inquiry for Scientific Thinking, Analytics, and Reasoning* (iSTAR) [27]. This tool expands on Lawson’s Classroom Test of Scientific Reasoning [22,28] and measures across skills commonly needed for students to systematically conduct scientific inquiry, which includes exploring a problem, formulating and testing hypotheses under controlled conditions, manipulating and isolating variables, and observing and evaluating the consequences. In order to demonstrate possible impact of specific pre-lab and in-class activities on student skill development, since 2016 we have administered iSTAR to all lab sections as a pre-test the first week of the term, and then the post-test was randomly assigned to a subset of the sections at weeks 4, 7, 9, 11, or 12 (end of course); these matched up with when lab reports were due indicating that an investigation cycle had been completed. In our analysis, the subgroups were treated as equivalent samples of the same student population due to the random assignment. In this way, by combining the multiple measures of different weeks, we were able to study the co-variations between students’ measured sub-skills and the learning activities that occurred during the course with a time resolution of 2-3 weeks (one inquiry cycle). Based on our results, we determined that most of the impact on student skill development was due to the first 4 weeks of the course, with little growth thereafter as shown in Figure 3.

As a result of this outcome, we made adjustments to the curriculum, with a particular focus on the material after week 4. One change was the inclusion of additional pre-lab activities that provide the students with practice in controlling variables in more complex multi-variable hypothetical scenarios, as well as additional practice in identifying situations which involve correlational versus causal relationships. Another addition to the pre-lab materials was the inclusion of articles found in the popular
press that include studies around multi-variable situations. Here, students are asked to evaluate the reported evidence and claim, thereby connecting the skills learned in class to a real-world event. These changes were implemented in Spring semester 2019 and assessment data indicates that student reasoning abilities continued to improve through the end of the course as shown in Figure 4.

![Figure 3](image3.png)

**Figure 3.** Student mean scores for iSTAR for Fall 2016 through Spring 2018, where n=2161 students in total.

![Figure 4](image4.png)

**Figure 4.** Student mean scores for iSTAR for Spring 2019, where n=558 students total.

In order to address the second research area regarding student abilities to engage in TEC, we looked at lab report writing scores over time. The ability to effectively communicate outcomes is an important 21st Century skill. The data in Figure 5 shows the scores of students’ lab reports for four consecutive lab investigations, which span the entire lab course (Pendulum, Newton’s 2nd Law, Simple Harmonic Motion, and Rotational Inertia). The data are mean scores of all students’ lab reports collected in Fall 2012, 2015, 2018, and Spring 2019. The data from Fall 2012 indicates that students struggled in writing reports that met our expectations. Between 2012 and 2015, additional support was added to the course materials to help report writing; including providing students with detailed grading rubrics and sample lab reports to guide their writing.

Due to this success, in 2018 we increased our expectations for students’ lab reports, and students were provided with a rubric that now included items for ability to engage in evidence-based reasoning for TEC. Question prompts were also added to the in-class instructions to support this new expectation, particularly in terms of generating evidence to support claims made. For example, students were now asked to critically examine the data and claims of at least two other lab groups in their class and comment...
on whether the work of these other groups supported or refuted their own claims. The overall pattern of scores in Figure 5 indicates that although the expectations for the lab reports had increased, students were successful in their lab report writing beginning with the first report (Pendulum).

In Spring 2019, we made additional adjustments to continue to support our students. The research literature indicates multiple areas for which students struggle with TEC, many of which were discussed earlier in this paper. We used these ideas to develop guiding reflection questions that students now address at the end of each inquiry cycle in our lab curriculum to better support this thinking process. In particular, these question prompts were designed to guide student thinking around the validity of the evidence gathered. For example, one of the prompts asks “How confident are you in the values obtained for each measurement? Be sure to comment on the ranges of uncertainty (variation) for these measurements.” Although students continue to be asked to critically examine the data and claims of at least two other groups, they are now directed to extend their ideas to consider other possible factors not tested by any of the groups and generate new hypotheses, if warranted. In addition, students are explicitly asked to compare their experimental model to the theoretical model and not only identify similarities, but also reconcile any differences. Figure 5 indicates that the changes made in 2019 further support students in their ability to generate evidence-based claims and engage in TEC in their writing. However, the scores do not continue to improve across the term, and preliminary data suggests that most of the difficulties are related to student abilities to connect their experimental models with scientifically accepted models, which is a core step in the TEC process for knowledge development (see Figure 1). This finding will guide future revisions to the lab curriculum as we continue to improve its impact on students.

![Figure 5](image.png)

**Figure 5.** Mean lab report scores for four lab investigations, with data collected across four semesters. Note that beginning in FS18 the lab report requirements increased to emphasize TEC.

6. **Conclusion**

This paper described the curricular framework that guided the development and evaluation of an introductory college physics lab curriculum designed to promote student abilities for engaging in the process of theory-evidence coordination. The research findings regarding student development of the subskills which support TEC, as well as their abilities to communicate evidence-based arguments in their lab report writing, indicates that the lab curriculum does in fact impact student abilities in these areas. However, the results also suggest that there is room for additional skill development. The lab
course activities will continue to be revised based on these outcomes. This is important as few college courses promote these skills, which are essential for the types of reasoning and decision-making skills needed for college level STEM learning in 21st Century education. In addition, the reasoning subskills that support TEC, including the ability to control variables, engage in data analytics, and causal decision-making, are foundational to a broader spectrum of scientific reasoning skills essential for knowledge acquisition and conceptual change [29], developing deep scientific understanding and conducting scientific inquiry [25], and student capabilities in general decision-making and problem solving [30].

Acknowledgment
The article is based on work supported by the National Science Foundation (NSF) under Grant No. DUE-1431908. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

7. References
[1] United States Chamber of Commerce 2017 Bridging the Soft Skills Gap: How the Business and Education Sectors are Partnering to Prepare Students for the 21st Century Workforce (Washington DC: Center for Education and Workforce, US Chamber of Commerce Foundation)
[2] National Science & Technology Council 2018 Charting a Course for Success: America’s Strategy for STEM Education (Washington, DC: Office of Science and Technology Policy)
[3] National Research Council 2012 Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century (Washington, DC: National Academies Press)
[4] Myatt, M 2012 6 Tips for making better decisions Forbes Leadership Column March 28
[5] Halpern, D F (1999) Teaching for critical thinking: Helping college students develop the skills and dispositions of a critical thinker New Directions for Teaching and Learning, 80 69-74
[6] Ennis, R 1993 Critical Thinking Assessment Theory into Practice 32 179-186
[7] Paul, R 1990 Critical Thinking: What Every Person Needs to Survive in a Rapidly Changing World (Rohnert Park, California: Center for Critical Thinking and Moral Critique)
[8] Kuhn, D 2004 Blackwell handbook of childhood cognitive development (Malden, MA: Blackwell) p 371–393
[9] Zimmerman, C 2005 The development of scientific reasoning: What psychologists contribute to an understanding of elementary science learning Retrieved from National Research Council’s Board of Science Education website www7nationalacademiesorg/bose/Corinne_Zimmerman_Final_Paperpdf
[10] Bao, L et al. 2009 Learning and scientific reasoning Science 323 586-587
[11] Zhou, S, Han, J, Koenig, K, Raplinger, A, Pi, Y, Li, D, Xiao, H, Fu, Z, and Bao, L 2016 Assessment of Student Reasoning: The effects of task context, data, and design on student reasoning in control of variables Thinking Skills and Creativity 19 175-187
[12] Kuhn, D 2007 Reasoning about multiple variables: control of variables is not the only challenge Sci. Educ. 915 710-726
[13] Reif, F, and Larkin, J H 1991 Cognition in scientific and everyday domains: Comparisons and learning implications J. Res. Sci. Teach. 28 733–760
[14] Stanovich, KE, and West, RF 1997 Reasoning independently of prior belief and individual differences in actively open-minded thinking J. Educ. Psych. 89 342–357
[15] Norman, O 1997 Investigating the nature of formal reasoning in chemistry: Testing Lawson’s multiple hypothesis theory J. Res. Sci. Teach. 34 1067–1081
[16] Lawson, A E 1992 The development of reasoning among college biology students - a review of research J. College Sci. Teach. 21 338-344
[17] Zeineddin, A and F Abd-El-Khalick 2010 Scientific reasoning and epistemological commitments: Coordination of theory and evidence among college science students J. Res. Sci. Teach. 47 1064-1093
[18] Kuhn, D 1992 Thinking as argument *Harvard Educ. Rev.* 62 155-178
[19] Klaczynski, PA, and Narasimham, G 1998 Development of scientific reasoning biases: Cognitive versus egoprotective explanations *Dev. Psych.* 34 175–187
[20] Ding, L, Wei, X, and Mollohan, K 2016 Does higher education improve student scientific reasoning skills? *Int. J. Sci. and Math. Educ.* 14 619-634
[21] Kuhn, D, Ramsey, S, and Arvidsson, T S 2015 Developing multivariable thinkers *Cog. Dev.* 35 92-110
[22] Lawson, A E 2000 Classroom test of scientific reasoning: Multiple choice version, based on Lawson, A E 1978 Development and validation of the classroom test of formal reasoning *J. Res. Sci. Teach.* 15 11-24
[23] Koenig, K, Schen, M, Edwards, M and Bao, L 2012 Addressing STEM retention through a scientific thought and methods course *J. College Sci. Teach.* 41 23-29
[24] Koenig, K, Schen, M, and Bao, L 2012 Explicitly targeting pre-service teacher scientific reasoning abilities and understanding of nature of science through an introductory science course *Sci. Educ.* 21 1-9
[25] Koenig, K, Wood, K, Bortner, L, and Bao, L 2019 Modifying traditional labs to target scientific reasoning *J. College Sci. Teach.* 48 28-35
[26] Kuhn, D 1989 Children and adults as intuitive scientists *Psych. Rev.* 96 674–689
[27] Bao, L iSTAR Assessment: Inquiry for scientific thinking and reasoning istarassessment.org
[28] Bao, L, Xiao, Y, Koenig, K, and Han, J 2018 Validity evaluation of the Lawson classroom test of scientific reasoning *Phys. Rev. Phys. Educ. Res.* 142 020106
[29] Bransford, J D, Brown, A L, and Cocking, R 2000 How people learn: Brain, mind, experience, and school (Washington, DC: National Academy Press)
[30] Williams, W M, Papierno, P B, Makel, M C, and Ceci, S J 2004 Thinking like a scientist about real-world problems: The Cornell Institute for Research on Children Science Education Program *Appl. Dev. Psych.* 25 107–126