

Combination of interventions can change students' epistemological beliefs

Calvin S. Kalman,^{1,*} Mandana Sobhanzadeh,² Robert Thompson,³
Ahmed Ibrahim,⁴ and Xihui Wang⁵

¹*Department of Physics, Concordia University, Montreal, Quebec H3G 1M6, Canada*

²*Department of Chemistry and Physics, Mount Royal University, Calgary, Alberta T3E 6K6, Canada*

³*Department of Physics and Astronomy, University of Calgary, Calgary, Alberta T2N 1N4, Canada*

⁴*Department of Psychology, University of California, Riverside, Riverside, California 92521, USA*

⁵*Department of Educational and Counselling Psychology, McGill University,
Montreal, Quebec H3A 0G4, Canada*

(Received 28 May 2015; published 21 December 2015)

This study was based on the hypothesis that students' epistemological beliefs could become more expertlike with a combination of appropriate instructional activities: (i) preclass reading with metacognitive reflection, and (ii) in-class active learning that produces cognitive dissonance. This hypothesis was tested through a five-year study involving close to 1000 students at two institutions, in four physics courses. Using an experimental design, data from student interviews, writing product assessments, and the Discipline-Focused Epistemological Beliefs Questionnaire (DFEBQ) we demonstrate that the beliefs of novice science learners became more expertlike on 2 of the 4 DFEBQ factors. We conclude that a combination of an activity that gets students to examine textual material metacognitively (Reflective Writing) with one or more types of in-class active learning interventions can promote positive change in students' epistemological beliefs.

DOI: [10.1103/PhysRevSTPER.11.020136](https://doi.org/10.1103/PhysRevSTPER.11.020136)

PACS numbers: 01.40.Fk, 01.40.gb

I. INTRODUCTION, MOTIVATIONS, AND GOALS

Over the past several decades, physics education research (PER) has developed many new pedagogical approaches (e.g., Peer Instruction [1], Physics by Inquiry [2], and Studio Physics [3]) and demonstrated their effectiveness through a range of tools such as concept-based pretests and post-tests (e.g., Force Concept Inventory [4] and the Brief Electricity and Magnetism Assessment [5]). As the field has achieved this level of maturity, this work, and the goals of our group as a whole, has shifted our focus beyond the question of whether or not a pedagogical technique is effective, towards more of a focus on answering the question of why a particular technique or class of techniques is effective. In attempting to answer these why questions, we look to established psychological and developmental concepts, theories, and models (e.g., cognitive dissonance [6] and epistemic development [7]) to both provide structure to our studies and support the analysis and interpretation of the results.

In particular, the work presented in this paper focuses on investigations of a set of pedagogical tools that attempt to improve student engagement with the course material outside the large lecture components of first-year physics courses through the use of targeted work-at-home

pedagogical tools (Reflective Writing [8] and argumentative essays [9]) and small group pedagogical tools (Conceptual Conflict Collaborative Groups [10,11], and laboratories [12]). The former tools focus on enabling the students to consider course materials metacognitively, while the latter are designed to produce cognitive dissonance in the students' consideration of the material. We postulate that if students reflect metacognitively on textual material before coming to class and then have interventions in class that have them examine subjects that produce cognitive dissonance, the students' epistemological beliefs would evolve from those characterizing a novice learning towards those consistent with a more expert learner. While it is relatively straightforward to measure students' understanding of physics content, it is much more difficult to measure what they believe about the nature of physics and learning physics (i.e., epistemological beliefs). Therefore, this work uses experimental and control student groups to test our postulate using both student interviews and writing product analysis (qualitative data) as well as the Discipline-Focused Epistemological Beliefs Questionnaire (DFEBQ) as a pre- and post-test (quantitative data), as presented and analyzed in the remainder of this paper.

In order to present this work clearly and completely, we first provide an outline of the relevant background theory and literature related to the topics of cognitive dissonances, as well as epistemological theory and models (Sec. II). We then outline our studies, including providing details on the pedagogical tools and experimental measures as well as the study groups (Secs. III and IV), before reviewing and

*Calvin.kalman@concordia.ca

analyzing the experimental results and conclusions (Secs. V and VI).

II. BACKGROUND THEORY AND LITERATURE REVIEW

A. Cognitive dissonance

Festinger [6] wrote, “In the course of our lives we have all accumulated a large number of expectations about what things go together and what things do not. When such an expectation is not fulfilled, dissonance occurs.” “He can even distort his perception and his information about the world around him. Changes in items of information that produce or restore consistency are referred to as dissonance-reducing changes” [6].

This is precisely the situation of the typical student in the gateway physics course in mechanics. Students have been experimenting since they were very young and their understanding of nature clashes with the Newtonian synthesis presented in the classroom and in their textbook. At the same time they have strong beliefs that knowledge is conveyed by authorities (instructor and textbook). This results in cognitive dissonance. To reduce the dissonance between their understanding and what they hear in the classroom and read in the textbook, students mishear the teacher and misread the textbook. Every time that we have given a seminar and mention about students coming up after class and stating that the instructor has said exactly the opposite of what the instructor said everyone in the room nods their head.

In our opinion cognitive dissonance might occur because of students epistemic beliefs. McCaskey stated [13].

If a student believes that knowledge in physics should come from a teacher or authority figure, and the class activities require more independent thought than direct intervention, there is epistemological conflict. Likewise, if a student comes in thinking that physics consists of a bunch of equations to be memorized, and the instructor focuses more on concepts, there is conflict. Finally, if a student is being presented material in a fragmented way, but he or she would expect or believe the material should fit together more cohesively, that would cause another type of conflict. These conflicts (or, conversely, a lack of these conflicts) can affect learning above and beyond specific difficulties with mathematics or concepts.

B. Application to physics education

“In typical physics classes, students’ beliefs deteriorate or at best stay the same. There are a few types of interventions, including an explicit focus on model building and/or developing expertlike beliefs that lead to significant improvements in beliefs” [14]. In traditional lecture courses as well as courses that use interactive engagement and lead to large gains on the Force Concept Inventory (FCI) [15]

and other content surveys, students’ scores on the Maryland Physics Expectations Survey (MPEX) [16] and the Colorado Learning Attitudes about Science Survey (CLASS) [16] at the end of the course are *less* expertlike than they were at the beginning. These courses usually have large enrolments, and are calculus based [14].

We conjecture that if students reflect metacognitively on textual material before coming to class and then have interventions in class that have them examine subjects that produce cognitive dissonance, the students’ epistemological beliefs would become more expertlike. The normal alternative to a lecture-based classroom format is a single type of intervention. We have been interested in seeing if an activity that gets students to examine textual material metacognitively (reflective writing) combined with one or more interactive interventions could help students change their approach to learning [17]. In this spirit, we investigated the effects of these interventions on students’ epistemological beliefs. Epistemological beliefs refer to individuals’ implicit thinking about the nature of knowledge and knowing, which have direct and indirect effects on student learning, such as the types of learning strategies that students use, their readiness for conceptual change and their self-regulated learning [18].

In pursuit of improved learning by students in university courses, one must first define how we measure the quality of learning. There are many diverse approaches to such a measurement including the following: (i) direct measurement of student knowledge through testing, including standardized testing and conceptual testing; (ii) measurement of student engagement as a proxy for learning as there exists a well established link between engagement and learning effectiveness [19–21]; or (iii) evaluating students’ approaches to learning and knowledge, specifically on the evolution from novice learning to expert learning [22–26]. Each of these approaches has its strengths and weaknesses in terms of accuracy of measurement and insight to explanations for effectiveness or lack thereof. This work focuses on the third of these approaches, as measurement of students’ epistemic thinking offers insight into both learning effectiveness and explanation for the maturation or lack of maturation in learning development.

C. Stages in epistemic development in students

In the 1950s and 1960s Perry [27] conducted a longitudinal study on college students’ understanding of the nature of knowledge involving lengthy interviews of students from various classes at Harvard and Radcliffe. In this study Perry demonstrated that college students progress through a series of nine stages corresponding to the way the students view the world in general, and their courses in particular. The first five stages correspond mostly with epistemology and intellectual development, whereas the last four stages deal with ethical and identity development. Perry’s model describes students’

epistemological responses to the college environment. The nine stages are descriptive of the nature of knowledge and truth that are categorized as follows: dualism, multiplicity, contextual relativism, and commitment within relativism. Perry suggests that how college students respond to the college environment is an evolving developmental process, which is brought about through cognitive disequilibrium. "Individuals interact with the environment and respond to new experiences by either assimilating to existing cognitive frameworks or accommodating the framework itself." [7], p. 91.

Perry's male sample brought it under attack in the late 1970s. Belenky *et al.* [28] decided to study women's ways of knowing. Though this model is different from Perry's model by focusing more on the source of knowledge and truth, it can be lined up to Perry's positions.

In 1986, Magolda [29] started a longitudinal study on college students. The sample included both male and female. Her epistemological reflection model contains four different "ways of learning." The research did not directly probe students' perspective of knowledge itself, but focused more on the nature of learning in the college classroom, with each way of learning corresponding to particular epistemic assumptions. With four ways of learning, we have four epistemic assumptions: absolute, transitional, independent, and contextual. Absolute knowers believe absolute knowledge that comes from authority; transitional knowers start to doubt the certainty of knowledge and authority; independent knowers begin to value their own opinions, and contextual knowers begin to construct their own perspective.

King and Kitchener's reflective judgment model [30,31] concerns students' epistemic assumptions behind reasoning. The model has seven reflective judgment stages that are categorized at three levels: prereflective, quasireflective, and reflective. Prereflective is parallel to Perry's dualism and quasireflective is parallel to Perry's multiplicity and contextual relativism. The highest level of reflection is similar to Perry's commitment within relativism. Throughout the seven stages, the focus is on "both the individual's conception of the nature of knowledge and the nature or process of justification for knowledge."

Kuhn's research model [32,33] of argumentative reasoning focuses on individual responses to everyday life. Through interviews on how college students reason and make judgments in everyday life, Kuhn studied the evidence for their epistemological thoughts. She reported three categories of epistemological views: absolutist, multiplist, and evaluativist that are aligned with Perry's four categories.

Though all of these models have different foci, we can see the similarity across the models: students progress through stages where they experience more and more uncertainty, and simultaneously, their way of acquiring knowledge changes from being passive to being more

active and constructive. Epistemological growth or development is a desired educational outcome that we educators should help students to achieve.

As different dimensions in those one-dimensional models are not necessarily developing at the same rate [34], it is valuable to develop multiple dimensional models. Concerning the number of dimensions, no agreement has been reached. We have three-dimensional models [35], four-dimensional models [7,34], and even indefinite numbers [36].

Schommer's model [34] includes four dimensions. Two of them are about students' way of learning: *fixed ability* and *quick learning*. The other two, *simple knowledge* and *certain knowledge*, are about the nature of knowledge.

Based on a review of all one-dimensional models and Schommer's four-dimensional model, Hofer and Pintrich [7] proposed two areas of epistemological theories: the nature of knowledge and nature of knowing, each having two dimensions. For the nature of knowledge, these dimensions are *certainty of knowledge* and *simplicity of knowledge*. For the nature of knowing, these are *source of knowledge* and *justification for knowing*. Each dimension is a continuum: *Certainty of knowledge* can take on values from fixed to tentative and evolving; *simplicity of knowledge* from discrete and concrete to relative and contextual; *source of knowledge* from transmitted by external authority to constructed during interaction with others by the knower; *justification for knowing* from accepting knowledge claims without evaluation to justifying claims with self-reasoning.

In comparing all of the one-dimensional models with Hofer and Pintrich's four-dimensional model, we find some components that are important yet are not included in the multiple dimensional model. For this reason, besides the four core dimensions, Hofer and Pintrich also proposed "peripheral dimensions": beliefs about learning, instruction, and intelligence.

Although there are many models and theories related to students' epistemological beliefs and learning, there is little work to test the dimensions suggested for science students. Hofer [37] is one of the researchers who showed that epistemological beliefs vary based on the field of study and tested the dimensions of personal epistemology of science students. In this study, we use Hofer's DFEBQ, which is adapted for physics to evaluate whether a combination of Reflective Writing and laboratorials change students' epistemological beliefs.

The discipline-focused questionnaire contains items that Hofer and Pintrich [7] adapted from existing instruments (Perry's Checklist of Educational Values and Schommer's epistemological beliefs questionnaire). They added more items to address the four discussed dimensions of epistemological models. A team of researchers familiar with the literature developed the questionnaire and three psychologists reviewed the questionnaire to check wording, content validity, and the relevance of the questions to each of

the four dimensions [37]. In this questionnaire each item typically refers to the field or subject matter as the frame of reference. Students respond to each of the 27 items on a five-point scale (1 = strongly disagree; 5 = strongly agree). Hofer used the DFEBQ to find the loadings of each item and identified four factors: (i) certain or simple knowledge (eight items), (ii) justification for knowing personal (four items), (iii) source of knowledge authority (four items), and (iv) attainability of truth (two items). In this specific analysis, certainty of knowledge and simplicity of knowledge emerged as one factor. “Justification for knowing personal” represents the view that knowledge and knowing are justified by personal opinion or firsthand experience. This factor does not contain questions related to evaluation of evidence or assessment of expert opinion. “Source of knowledge authority” is about the extent to which students rely on expert knowledge, texts, and any external authority as the source of knowledge. An additional factor called “attainability of truth” showed up in Hofer’s analysis regarding the degree to which students believe that truth is attainable by experts. Given the extensive validation procedure used by Hofer and its relation to many previous extensively tested previous efforts beginning with Perry and the existence of a form designed for physics, the DFEBQ was selected for this study.

Ideally, we would have our students respond to the DFEBQ and do a factor analysis of our results. We, however, did not feel that we had a sample big enough for this purpose. Instead, we decided to use the factors found by Hofer. This is a possible limitation in this study.

1. The dimensions of personal epistemology in Hofer’s questionnaire

As discussed in the previous section, Hofer [37] believes that there are two main dimensions that are common among the various epistemological models: the nature of knowledge (what one believes knowledge is) and the nature or process of knowing (how one comes to know). There are “certainty of knowledge” and “simplicity of knowledge” dimensions under the area of nature of knowledge. In the DFEBQ these two dimensions emerge as a single factor. The area of nature of knowing contains “source of knowledge” and “justification of knowledge” dimensions. Additionally, as noted above, an additional factor called “attainability of truth” showed up in Hofer’s analysis regarding the degree to which students believe that truth is attainable by experts. In our analysis we use the four factors: certainty or simplicity, source of knowledge, justification for knowing, and attainability of truth found by Hofer.

Certainty of knowledge.—This factor is related to the extent to which one considers knowledge as fixed and solid or more fluid and dynamic. The developmental psychologists consider this as a continuum that changes over time from a fixed to a more fluid view [37]. At lower levels, there is an absolute truth with certainty, while at higher levels

knowledge is considered uncertain and evolving. King and Kitchener [30] believed that students are open to new interpretation at the highest stage of reflective judgment. Kuhn [32] argued that at the highest stage the individuals welcome the possibilities and challenges to modify their ideas. Borda [38] believed that the way introductory science courses are presented leads students to consider science as a body of fixed facts.

Simplicity of knowledge.—As discussed earlier, Schommer [34] believed that knowledge is viewed on a continuum as a body of facts or as highly interrelated concepts. We also reviewed the other models considering the lower level view of knowledge as discrete, concrete, knowable facts and the higher level view of knowledge as relative, contingent, and contextual. We have designed many questions in the introductory physics laboratories to help students think about the concepts discussed previously and relate them to the new concepts taught. In the Reflective Writing activity, relating recently introduced key concepts to previously studied concepts is one of the main marking criteria. Therefore, this factor is of great interest to study in this research project.

Source of knowledge.—At lower levels of most of the epistemological models, knowledge is created outside the individual and exists in external authority. Refs. [27,37] described this shift in students’ epistemology as being a holder of meaning to a maker of meaning and being a part of the learning procedure. Similarly, King and Kitchener [30] described a shift in the process of learning at the highest stages from being an observer and receiver of knowledge to an active constructor of meaning.

Justification for knowing.—This factor is related to evaluation of knowledge and the extent to which they use evidence or authority to justify knowledge claims, and explore whether they use evidence or rely on authority and expertise. In the lower levels of the reflective judgment model [30], students justify knowledge through observation or authority. They may also justify knowledge based of what feels right. While at higher levels students use reasoning and begin to personally evaluate knowledge and the views of experts. It is interesting to explore this factor since many students try to use experimental results to prove what they see in the textbook or what they learned in the classroom. Students often explain experimental results based on what feels right. We indeed have faced students who tried to change the gained experimental results to be able to prove what they were taught in the course.

Attainability of truth.—Finally, the factor *attainability of truth* reflects individuals’ beliefs about whether ultimate truth is attainable. At one end of the continuum of truth being obtainable, individuals believe that people can

ultimately find the correct answer to any question. At the other end of the continuum, wherein ultimate truth is not attainable, individuals believe that not every question has a correct answer.

III. RESEARCH PLAN, HYPOTHESIS, AND STRUCTURE

Saul Finster's discussion of the epistemology of chemistry [39,40] was one of the first attempts to apply Perry's model in the sciences. Finster pointed out that understanding students' epistemologies allows teachers to create better environments for promoting student learning. In the same vein, Bendixen, Dunkle, and Schraw [41] have demonstrated that the quality of student learning is affected by students' epistemological beliefs. Elby [42] has noted that "students' epistemological beliefs—their view about the nature of knowledge and learning—affect their mindset, metacognitive practice, and study habits in a physics course" (p. S64). Similarly, it has been argued that students' acceptance and understanding of evolution theory depends on their beliefs about the nature of science as well as their beliefs about the nature of knowledge [43–45].

Smith and Wenk [46] note "Within the science education community, there have been debates about the amount of coherence that exists in students' epistemological thinking, between those with knowledge-in-pieces viewpoints versus intuitive theories viewpoints" (p. 775). Smith and Wenk present data that they believe argue for "some underlying coherence in students' epistemological thinking." However, in contrast to their findings, a number of studies have shown that many students view science as weakly connected pieces of information to be separately learned, in contrast to the web of interconnections perceived by their instructors [47–50]. Support for this view is found in a study by Huffman and Heller [51] of 750 university students in a calculus-based introductory physics course that shows that most students' personal (alternative) scientific conceptions "are best characterized as loosely organized, ill-defined bits and pieces of knowledge that are dependent upon the specific circumstances in question" (p. 141). According to Paul Hewitt [52], students' actual epistemology is quite different from that of their teachers; while professors classify problems in terms of physics concepts, students tend to classify them by situations. Assuming this to be true, developing a scientific mindset may involve more than a conceptual change from personal scientific concepts to scientifically accepted concepts. It may also require a change in attitude; from a view that study in science is a matter of solving problems using an independent set of tools, classified according to problem type, to a view that a science subject consists of a web of interconnected concepts.

Havdala and Ashkenazi [53] (p. 1137) contrast the two main paradigmatic approaches to the study of personal epistemology: epistemological beliefs [54] and

epistemological development [27,29,30,33]. According to Schommer [34], personal epistemology would be better portrayed as a system of more or less independent beliefs in which individuals may be sophisticated in some beliefs, but not necessarily sophisticated in other beliefs. In contrast, epistemological development suggests that individuals move through a one-dimensional developmental sequence—from a naive belief that knowledge is certain and directly accessible, to a mature view of knowledge being justified by integrating and evaluating different opinions and multiple sources of data. There is no real contradiction. As Finster [40] noted, Perry identified three methods by which students may interrupt the normal progression: temporizing, escaping, and retreating, all of which represent delays or temporary regression in the scheme. In this manner, Finster notes an important caveat about students' epistemological development at university. Students may utilize a relativist position in one class and a dualist in another. Indeed, the format of the class may influence the manner in which a student exercises a personal epistemology. Watters and Watters [55] point out that "information delivered in a mass lecture format accompanied by multiple choice examinations establishes a context that fosters learning by memorization." (p. 20).

This work aims to investigate our conjecture of the positive impact of prelecture metacognitive study on students' epistemic thinking through a pair of investigations at different institutions using a range of pedagogical tools and a range of measurement tools to determine effectiveness. In the following section, we will describe the pedagogical tools, and the measurement tools employed in this work. In Sec. V we will describe the two studies including goals, tools, and results. Section VI will conclude and bring together the results of the two studies.

IV. PEDAGOGICAL TOOLS AND EFFECTIVENESS EVALUATION INSTRUMENTS

In order to undertake a study of students' epistemic thinking, we must define the learning techniques that we are studying and, separately, outline the instruments or measures that we intend to use, and the benchmarks that will define success, in order to determine the effectiveness of our pedagogical techniques. Section IVA will outline the pedagogical techniques that are utilized in this work, while Sec. IV B will outline our measures.

A. Pedagogical tools

Our studies were designed to study and understand the effectiveness of using Reflective Writing in combination with various other pedagogical tools, including Conceptual Conflict Collaborative Groups, argumentative essays, and laboratories. The following subsections provide brief introductions to each of these four pedagogical tools. Section V

will outline which tools were used for each study, and why the particular combinations were selected.

1. Pedagogical tool: Reflective writing (RW)

Reflective writing [8] is a metacognitive activity, which has students examine textual material before coming to the classroom in the manner of a hermeneutic circle [56]. The hermeneutic approach starts by having students initiate a self-dialogue about each textual extract. Within the framework of such a dialogue, there exist two “horizons.” There is the horizon that contains everything that a student believes from the particular vantage point of encountering the textual extract. The second horizon encompasses the potential in the textual extract; the sense in which the words, in the textual extract, are related within the language game understood by the author of the textbook. The student approaches the textual extract with preconceptions (misconceptions) about the material within the textual extract. The key quintessential experience occurs when the student is pulled up short by the textual extract. “Either it does not yield any meaning or its meaning is not compatible with what we had expected” [56], p. 237. At this point the dialogue begins. The student questions what is known within the entire horizon [8], p. 163.

2. Pedagogical tool: Conceptual Conflict Collaborative Groups (C³G)

The in-class Conceptual Conflict Collaborative Group exercises [10,11] were designed to provide students a learning environment to question their alternative personal scientific conceptions and to expose them to other perspectives. For example, in one of the exercises, students were asked to compare the motion of a free-falling body with a horizontal projectile. Then two groups of students were asked to present their ideas and to have other students question and challenge their proposed ideas. Once students’ perspectives are exposed to “public” scrutiny, their certainty about knowledge is questioned or reevaluated.

3. Pedagogical tool: Argumentative essays (AE)

This critique writing activity [9] is basically an argumentative essay, in which students have to put forward as many possible arguments in favor of all the conceptual viewpoints raised in class and then point out which viewpoint is correct from an experimental point of view. Argumentation may promote deeper processing of the learning content, and students’ dispositions to engage in argument are closely related to the epistemic beliefs they hold [57], and will reciprocally influence their beliefs about knowledge.

4. Pedagogical tool: Laboratorials (Lbt)

Laboratorials were developed by the Physics Education Development group at the University of Calgary [12]

inspired by the introductory physics tutorial system in the University of Washington [58]. The curriculum used at the University of Washington is entitled “Tutorials in Introductory Physics” and was written by the Physics Education Group at the University of Washington. The tutorials are worksheets that require students to work through concepts that have been identified by research to be particularly difficult. Some require students to perform experiments. However, there is still a traditional laboratory system for the first year physics courses at the University of Washington.

“Laboratorial” (Lbt) comes from a combination of “laboratory” and “tutorial.” In this new style of lab, students use a worksheet with conceptual questions, calculation problems, and instructions for experiments and computer simulations. Lbt highlight physics concepts covered in lectures and encourage students to present and share their ideas with one another. Each Lbt worksheet starts with conceptual questions and then asks students to make predictions. After doing the experimental part, students need to explain whether their results support their prediction or not. Each lab section has one lab instructor assigned to a maximum of 16 students. In Lbt, students complete an Lbt worksheet in groups of 3 or 4 students. There are usually 3 to 6 checkpoints on each worksheet. The purpose of the checkpoints is to encourage an ongoing interaction between the students and lab instructor. Each time the students reach a checkpoint, they review the answers with the lab instructor. If the answer to a question is wrong or students are not proceeding in the right direction, the lab instructor leads the students to find the correct answer by themselves, exploring and discussing alternative ideas. The Lbt worksheets are prepared and tested such that students who arrive on time and concentrate on the material can finish all checkpoints in the time allotted.

B. Students’ epistemic thinking

In order to determine the effect of the combination of the above-listed pedagogical tools on students’ epistemic thinking, in particular, on student beliefs around their understanding of physics, we used a range of student surveys, interviews, and writing product assessments. In the case of this study, we used qualitative data sources (student interviews, analysis of student writing products) to investigate student beliefs and epistemic approach, but we focused our quantitative work on the results of one student survey instrument, the Discipline-Focused Epistemological Beliefs Questionnaire [37]. The measurement of epistemic belief has been considered a daunting task. We are aware that the validity of survey investigation in this area needs to be further examined and improved to make it more situated with the interested domain, context or tasks, and focusing on one survey alone constitutes a major limitation of the current study. Therefore, understanding of our quantitative

results was coupled with the above-mentioned qualitative data we have collected and analyzed. In addition, instruction was developed to prompt participants to respond to the survey questions by thinking in the domain of physics.

1. Measurement tools: Discipline-focused epistemological beliefs questionnaire

The Discipline-Focused Epistemological Beliefs Questionnaire [37] is a multi-question ($N = 27$) student completed survey that the students complete in paper form. The questions are categorized into four areas: certainty or simplicity (CS), justification of beliefs (JB), source of knowledge (SK), and attainability of truth (AT), with student scores on all questions in a given area combined to generate an overall score for each of the four categories. Therefore, the DFEBQ provides us with independent measures of four epistemic belief factors (CS, JB, SK, AT). The CS factor represents individuals' beliefs about the nature of knowledge. The certainty component of this factor reflects whether individuals view knowledge as being absolute and certain or as tentative and evolving, whereas the simplicity component reflects whether individuals believe knowledge is accumulated bits of facts or is interconnected and context specific. The JB factor reflects a belief that knowledge is justified by relying on experts versus individual opinion and firsthand experience. The SK factor reflects beliefs that knowledge is handed down by an authority figure such as a teacher or other expert, or that it can be personally constructed. Finally, the factor AT reflects individuals' beliefs about whether ultimate truth is attainable. At the one end of the continuum of truth being obtainable, individuals believe that people can ultimately figure out the correct answer to any question. At the other end of the continuum, wherein ultimate truth is not attainable, individuals believe that not every question has a correct answer.

2. Measurement tools: Semistructured interviews (SSI)

Semistructured interviews (SSI) with individual students were used to develop qualitative data to assess the impact of pedagogical tools used in this study. The interview questions and the rubrics were initially developed in the first year of the Montreal and BC project (see Sec. VA). In the second year we pilot tested the interview questions and rubrics. After refining the interview questions and the rubrics in the third year we conducted our research. A fourth year was taken for the analysis of the Montreal and BC results. The script was then reviewed and adapted slightly for the Calgary project (see Sec. VB). The interview questions focused on four main topics: students' perspectives on preunderstanding, the general way of learning this course, the main aspects that make Reflective Writing a successful activity, and students' perspectives on Lbt. Each interview consisted of 13 scripted questions, with additional follow-up questions, which varied from student to student. Each interview was completed in

approximately 1 h. Interviews for the Montreal and BC project were carried out by members of our research team who were not instructors for the individual students. All but one of the authors are not part of the institution where the Calgary experiment took place and were not involved with teaching the students. For the Calgary project interviews were carried out by the remaining student author, who taught 8 lab sections (4 sections were asked to prepare summary writing and 4 other sections that completed RW assignments). The other 4 RW and three summary writing sections were taught by two other instructors. We make a comparison between the groups taught by the author and those taught by the other instructors in the results section. Students in both groups were randomly mixed in four classroom sections. All instructors use clickers in the class. The marking criteria and assignments and activities are the same in all sections. The classroom instructors were not part of our research team.

3. Measurement tools: Student writing products (SWP)

The second source of qualitative data used in this work was a rubric-based analysis of the students' writing products. Rubrics were developed for the RW and AE writing products, and used in both the Montreal and BC and Calgary studies. The rubrics specifically targeted a determination of the students' epistemic approach, as reflected in the writing, and were used both as an independent measure and as a cross-comparator to confirm the accuracy of student self-assessment as expressed in their interviews. SWP analysis was carried out on the work of students who participated in SSI sessions, and on representative samples of students not involved in the interview process.

V. STUDY DESCRIPTIONS AND RESULTS

Over the last five years, a pair of studies were undertaken to better understand the effect of metacognitive exercises in physics on students' epistemic thinking. The two experiments, coined the Montreal and BC study and the Calgary study, are complementary in that they utilize different pedagogical tools, and different student sets and settings, but the same quantitative instrument (DFEBQ), similar qualitative tools (SSI and SWP), and a common goal.

A. The Montreal and BC study: Reflective Writing, collaborative group, and argumentative essay

1. Hypothesis

The Montreal and BC study over a period of three years (Table I) investigated the hypothesis that 1st year students of physics would experience a more effective learning environment and achieve strong improvement in their learning style (from novice to expert learner) through the use of a metacognitive pedagogical exercise, specifically RW, in combination with interactive pedagogical exercises, specifically C³G, and AE.

TABLE I. Overall design of the Montreal and Vancouver study across three years.

| Prestudy rudimentary interview questionnaire and Reflective Writing products | | | | | |
|--|--|---|--|----------------------------------|--|
| Institution A: University | | | Institution B: Community college | | |
| <i>n</i> | 75 Students | | 31 Students | | |
| Three year study | | | | | |
| Experimental groups | | | Control groups | | |
| | Institution A: University | Institution B: Community college | Institution A: University | Institution B: Community college | |
| Year 1 2010 activity | Developing interviews and rubrics for interviews, Reflective Writing, and critique writing | | Developing interviews and rubrics for interviews and summary writing | | |
| Pilot testing all instruments winter, spring 2011. | | | | | |
| Year 2 2011 <i>n</i> | 100 | 32 | 100 | 32 | |
| Year 3 2012 | Intervention | <ol style="list-style-type: none"> 1. Reflective Writing 2. Reflective write pair share and conceptual conflict group exercise 3. Critique writing | | none | |
| | Measures | <ol style="list-style-type: none"> 1. Pre- and postintervention interview rubric 2. Reflective Writing rubric 3. Critique writing rubric | | | |
| <i>n</i> | 100 | 64 | | | |

2. Background and methods

Participants in the Montreal and BC study were first-year students in two post-secondary institutions (Concordia University [Montreal] and Langara College [British Columbia]) enrolled in courses for nonphysics majors. At Concordia University, a university with a substantial graduate school, classes were relatively large sections (over 100 students each) of a typical calculus-based course in mechanics. At Langara College, a community college, there were relatively small classes (32 students each) of a typical algebra-based introductory course in mechanics, electricity, and magnetism. The two institutions used different textbooks and had different formats. However, at both institutions students were randomly assigned to all of the sections by the Registrar's office. The majority of students enrolled in both courses are science or engineering majors, with the courses being a requirement at the respective institutions. Generally, however, a small number of students from such cohorts will continue on in physics as their major. The reflective write-pair-share combined with C³G exercises activities were done during regular class hours, but did not count in the grades. The RW, summary writing, and AE were assigned as homework. Summary writing is a skill used by many students and for which some research is available showing positive effects of summary writing on recall and understanding [59]. The experimental group was exposed to all three instructional activities while the control group was only asked to perform SW of textual material before coming to class. All sections at one institution, experimental and control groups, were taught by the same instructor, who was not part of the research team that studied the students.

The purpose of this study was to investigate if and how the combined implementation of the entire suite of activities

could change students' approach to learning physics over and above the impact of each approach undertaken alone, and also if it could enhance their learning. To measure if the objectives were met we employed SSI at the beginning and the end of the semester and we examined the SWP produced by the students each week in the course. The SSI provides the primary data for this study. To confirm that students were actually doing what they said they were doing in the SSI, we triangulated the data garnered from the interviews with our analysis of the SWP.

The overall design of this experiment is shown in Table I. SSI data about how students think about concepts were collected in year 4 of the study, but certainly needed to be supported by data about student actions. We relied on a vast amount of data from analyzing SWP by means of rubrics, specifically from RW products done before they came to each class and from the AE written one to two weeks after particular classes. Altogether we triangulated the data garnered from the interviews with our analysis of the SWP. Overall, this study had potential access to data from 428 students.

The students gave evidence of how they viewed learning and how they changed their ways of learning as a result of the suite of activities. Student learning shifted toward more expertlike thinking, and demonstrated that easy-to-implement instructional innovations such as the suite of activities described in this study can narrow the novice-expert gap in the ways students engage physics as a subject.

Towards the end of this experiment a further experiment took place involving 40 students in an experimental section and 15 students in a control group, who elected to write the DFEBQ, administered at the beginning and end of the semester. The experiment and control section students differed by their learning experience, with only

experimental students participating in all of the RW, AE, and CCG activities.

3. Results Montreal and BC epistemological experiment

Our results based upon 426 students' responses indicated that students who experienced the full suite of activities become more expertlike after the one-semester intervention, beginning to see physics knowledge as interconnected and evolving, which can be better learned by relating the material to their prior knowledge and their life experience. Students who experienced summary writing did not experience such a change.

55 students took part in the epistemological experiment, which is insufficient to provide more than an indication that the suite of interventions would produce the desired result. We do not present the limited Montreal and BC DFEBQ results here. Instead we focus on the Calgary results where many students wrote the DFEBQ.

B. The Calgary study: Reflective Writing and laboratorials

1. Hypothesis

In this experiment, coined the Calgary study, we followed reflective writing with Lbt as an interactive intervention. The Calgary study investigated the hypothesis that 1st year physics students would experience a more effective learning environment and achieve strong improvement in their learning style (from novice to expert learner) through the use of a metacognitive pedagogical exercise, specifically RW, in combination with an interactive pedagogical exercise, Lbt.

2. Background and methods

At Mount Royal University the old style laboratory experimental exercises were changed to Lbt in introductory physics courses for first year students. These courses (Phys1201 and Phys1202) include 3 lecture hours per week, with weekly 2 hour tutorials (using Lbt) to cover the fundamentals of classical mechanics such as kinematics, Newton's laws of motion, and the concepts of work, energy, and linear momentum in the first semester and thermodynamics plus selected topics in electricity and magnetism in the second semester. Course materials and homework are available online using *Mastering Physics*, and students are expected to read the material before coming to class. Instructors make extensive use of "clickers" during the lecture.

Most students taking introductory physics courses are enrolled in the general science majors. At the start of the second academic year, students wishing to enter one of the other majors such as chemistry, geosciences, and biology will be asked to declare their intent officially to do so. All students enrolled as General Science majors have to take Phys1201 and Phys1202.

To find out whether any possible epistemological change is a result of using RW and Lbt, we designated two groups in fall 2014: an experimental group and a control group. Both groups performed laboratorials. The difference between the two groups was that the experimental group performed RW and the control group performed summary writing.

In fall 2014, in eight Phys1201 Lbt sections students were assigned to do RW, while in the remaining seven Phys1201 Lbt sections students were assigned to provide summary writing. Students were randomly assigned to all sections by the Registrar's office. There were 110 students in the experimental group and 102 students in the control group during the fall 2014 semester. All of these students took the epistemological survey.

To find out how students' epistemological beliefs change during two semesters, we gave 115 students from Phys1201, who enrolled in the Phys1202 course in winter 2015 an epistemological survey at the end of the semester. Students who had taken RW in Phys1201 continued with RW. Students who had taken summary writing in Phys1201 continued with summary writing. All students performed laboratorials in Phys1202.

The quantitative epistemological data address whether the combination of RW and Lbt have an impact on students' personal epistemologies, whereas the qualitative data focus on the key aspects that make RW a successful activity and students' perspectives on Lbt. In this mixed methodology, we have used an embedded design to provide additional sources of information not provided by the qualitative data.

In this study, we have addressed two main research questions:

- (a) What characteristics of reflective writing make this activity helpful?
- (b) What are students' perspectives on Lbt?

We collected and analyzed two kinds of qualitative data—students' reflective writing products and interview transcripts to address our first main research question. The last part of the interview focused on students' perspectives on Lbt. The broad question "What characteristics of RW make this activity helpful?" can be broken up into the following subquestions:

- (a) Does reflective writing change students' ways of learning physics?
- (b) Do students' writing products improve during the semester?
- (c) Do students have a hermeneutical approach while doing RW?
- (d) Does the combination of RW and introductory physics Lbt change students' epistemology?
- (e) How do students prepare their RW assignments?

Semistructured interviews were completed at the beginning and end of the 13-week course. We compared the pre- and post-interviews to see if there were reoccurring categories that might reveal underlying themes regarding students' views towards RW and Lbt. We also looked at

interviewee SWP to see if what they said was consistent with what they actually did during the semester and assessed the SWP of noninterviewed students to triangulate with and improve the results.

For the purposes of this paper, we also briefly looked at the question “Does RW change students’ ways of learning physics?” We asked students to explain how they studied for the Phys1201 course. Based on interviewees’ responses to this main question, we asked specific probe questions to find out what strategies they took to learn physics during the semester and how RW and Lbt influenced their learning strategies during the semester. Interviewees also talked about their expectations of the Phys1201 course and Lbt and the strategies they took to meet their expectations by the end of semester.

3. Results

We pursued an extension of the previous study with a larger sample (212 students; 110 in an experimental group

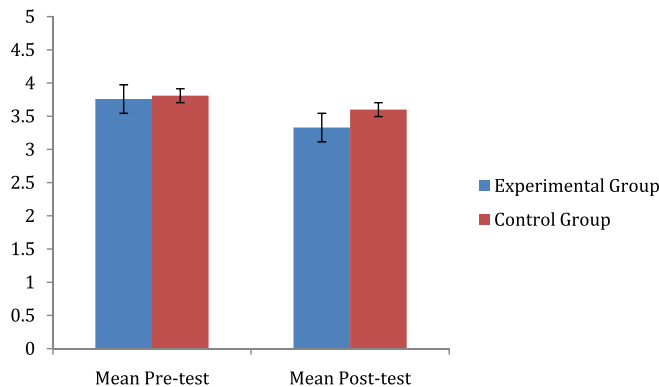


FIG. 1. Calgary study. Mean pre- and post-score of the *certainty or simplicity knowledge* factor for experimental and control groups with 5% error bars. (Low score is more expertlike reasoning.)

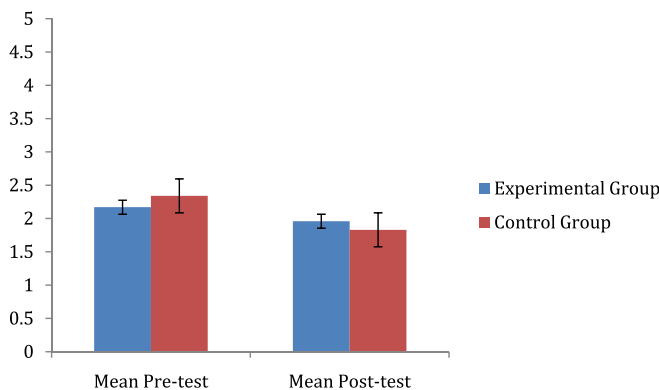


FIG. 2. Calgary study. Mean pre- and post-score of the *justification for knowing* factor for experimental and control groups with 5% error bars. (High score is more expertlike reasoning.)

and 102 in a control group). Based upon the interviews we evaluated the effect of Lbt on learning skills. We found that Lbt value students’ pre-understanding and that Lbt helps them integrate their preunderstanding with the experiments through thinking, comparison, reasoning, and explaining. They help students explore the relationship among various physics concepts. Our analysis indicated that laboratorials had helped students improve their understanding of the concepts.

We interviewed seven students who completed RW assignments. We also assessed the writing assignments of 30 random students to triangulate with and improve the results. We identified the specific aspects that appear to make reflective writing a successful learning activity. Our results indicate that reflective writing expands students’ possibilities of using preunderstanding and helps them expand their horizons. RW also affects students’ learning skills by improving their understanding of concepts, problem solving skills, engagement, and performance in the lab.

The comparison of the scores on the epistemological survey between the experimental group (students who did RW) and the control group (students who did summary

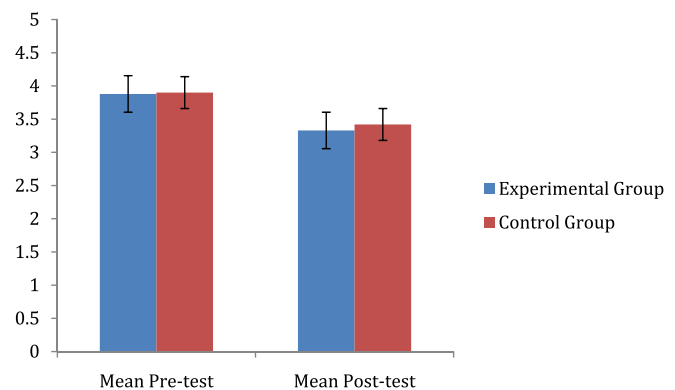


FIG. 3. Calgary study. Mean pre- and post-score of *source of knowledge authority* factor for experimental and control groups with 5% error bars. (Low score is more expertlike reasoning.)

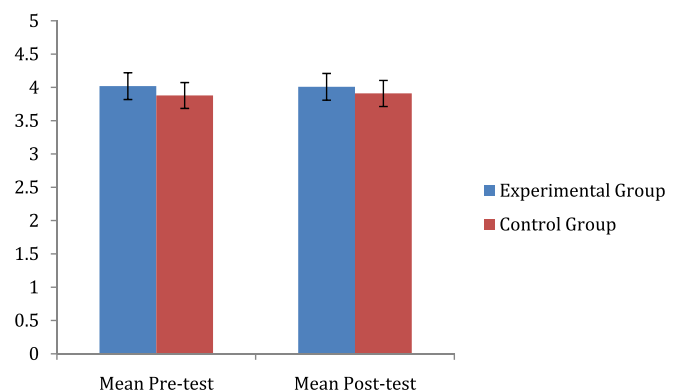


FIG. 4. Calgary study. Mean pre- and post-score of *attainability of truth* factor for experimental and control groups with 5% error bars. (Low score is more expertlike reasoning.)

TABLE II. The results of paired t test for the experimental group in the Calgary study.

| Factors | Mean pretest | Mean post-test | N | Paired differences (“Pretest”–“Post-test”) | | | t | Significant (2-tailed) |
|-------------------------|--------------|----------------|-----|--|--------------------|---------------------|------|------------------------|
| | | | | Mean | Standard deviation | Standard error mean | | |
| Certainty or simplicity | 3.76 | 3.33 | 110 | 0.436 | 0.290 | 0.0276 | 15.8 | 0.000 |
| Justification: personal | 2.17 | 1.96 | 110 | 0.214 | 0.386 | 0.0368 | 5.81 | 0.000 |
| Source: authority | 3.88 | 3.33 | 110 | 0.557 | 0.598 | 0.0570 | 9.77 | 0.000 |
| Attainability of truth | 4.02 | 4.01 | 110 | 0.014 | 0.644 | 0.0614 | 0.22 | 0.825 |

TABLE III. The results of paired t test for the control group in the Calgary study.

| Factors | Mean pretest | Mean post-test | N | Paired differences (“Pretest”–“Post-test”) | | | t | Significant (2-tailed) |
|-------------------------|--------------|----------------|-----|--|--------------------|---------------------|-------|------------------------|
| | | | | Mean | Standard deviation | Standard error mean | | |
| Certainty or simplicity | 3.81 | 3.60 | 102 | 0.218 | 0.344 | 0.0341 | 6.40 | 0.000 |
| Justification: personal | 2.34 | 1.83 | 102 | 0.510 | 0.745 | 0.0737 | 6.92 | 0.000 |
| Source: authority | 3.90 | 3.42 | 102 | 0.480 | 0.694 | 0.0687 | 7.00 | 0.000 |
| Attainability of truth | 3.88 | 3.91 | 102 | −0.025 | 0.794 | 0.0786 | −0.31 | 0.756 |

writing) are presented in Tables II and III and Figs. 1–4. The results of paired t tests for each factor are presented in Tables II and III. The data presented in Tables II and III are plotted as a bar graph in Figs. 1–4. Table IV shows that for the CS, SK, and AT factors a low score indicates more expert-type reasoning, while for the JB factor a high score points out more expertlike ways of thinking. The statistical test chosen to address this purpose was the paired t test [60]. The null hypothesis for our study states that there are no differences between the students’ responses to the DFEBQ at the beginning and end of the fall semester. The significance level (or p value) is used to measure the probability that the null hypothesis is true. We compared the experimental group with the control group to see whether the combination of labatorial and reflective writing activity had any effect in epistemological change of each factor. A one-way ANCOVA was conducted to determine a statistically significant difference between the experimental group and the control group on the post-test controlling for the pretest. Tables V–VII, show the statistical results of the ANCOVA test. Note that for the CS factor the effect size is

TABLE IV. The relationship between the score and the way of thinking in each factor.

| Factor | High score | Low score |
|---------------------------------|----------------------|----------------------|
| Certain or simplicity knowledge | Novicelike reasoning | Expertlike reasoning |
| Justification for knowing | Expertlike reasoning | Novicelike reasoning |
| Source of knowledge authority | Novicelike reasoning | Expertlike reasoning |
| Attainability of truth | Novicelike reasoning | Expertlike reasoning |

0.63. Most researchers consider that any effect size 0.5 or larger is an important finding.

Overall, the gain in the CS factor is significant. Figure 1 and Table V show that novice science learners become more expertlike after the one-semester intervention. Both control and experimental groups start off at the same level. Both groups become more expertlike but the experimental group makes more progress than the control group. As shown in Table V, this result is statistically significant.

There is a negative or regressive change in students’ epistemological beliefs for both control and experimental groups regarding the JB factor (Fig. 2). The control group starts off with more expertlike thinking than the experimental group. Both groups become less expertlike but the decrease for the experimental group is substantially less than that for the control group. As seen in Table VI, the difference between experimental and control groups is significant, which means the change in students’ epistemological beliefs on the JB factor for the experimental group is less than the epistemological change that the control group experienced. Loss in development on the JB factor due to cognitive dissonance is not unknown at this stage of students’ educational development, so minimizing loss in JB is a positive result.

Figure 3 shows an epistemological change in both the control group and the experimental group for the SK factor. Both control and experimental groups start off at the same level. However, as seen in Table VII the difference between experimental and control group is not statistically significant.

As shown in Fig. 4, there is no epistemological change in either control or experimental groups regarding AT. Therefore, the combination of Lbt and RW had no effect on AT.

In summary, on the CS and SK factors the experimental group became more expertlike. On the JB and AT factors,

TABLE V. A pairwise comparison of the variance of pre-post gains between the experimental group and the control group for the *certainty or simplicity of knowledge* factor.

| Group (I) | Group (J) | Mean difference (I-J) | Standard error | Significant | 95% Confidence interval for difference | | Effect size |
|-----------|-----------|-----------------------|----------------|-------------|--|-------------|-------------|
| | | | | | Lower bound | Upper bound | |
| Cont | Exp | 0.231 | 0.043 | 0.000 | 0.146 | 0.317 | 0.63 |

TABLE VI. Pairwise comparison of the variance of pre-post gains between experimental group and control group for *justification for knowing* factor.

| Group (I) | Group (J) | Mean difference (I-J) | Standard error | Significant | 95% Confidence interval for difference | | Effect size |
|-----------|-----------|-----------------------|----------------|-------------|--|-------------|-------------|
| | | | | | Lower bound | Upper bound | |
| Cont | Exp | -0.199 | 0.069 | 0.004 | -0.336 | -0.063 | 0.40 |

the experimental group had no significant change. On CS the control group became more expertlike but significantly less so than the experimental group. On JB the control group became significantly less expertlike. On SK the control group became more expertlike at the same rate as the experimental group. On AT the control group had no significant change.

We also examined whether the student epistemologies are stable and coherent. Many students who take the Phys1201 course in the fall take the Phys1202 course in the winter. In winter 2015, there were 63 students enrolled in Phys1202 who had been in the experimental group in Phys1201 in fall 2014 and completed reflective writing assignments in Phys1202 and took the epistemological survey. There were also 52 student enrolled in Phys1202 in winter 2015 who completed summary writing in fall 2014 and continued with summary writing in Phys1202 and took the epistemological survey. To find out how students'

epistemological beliefs change during two semesters, we gave these 115 students enrolled in the Phys1202 course in winter 2015 an epistemological survey at the end of the semester. In Tables VIII and IX the "mean pretest fall" refers to the fall 2014 scores at the beginning of the fall 2014 semester for the 115 students who were enrolled in Phys1202 in winter 2015. Note that since these students are a subset of those in Tables II and III, the scores for the beginning of the fall semester 2014 found in these tables do not exactly match the scores for beginning of the fall semester in Tables II and III.

We also compared the postscores of the epistemological survey in fall 2014 and the postscores of the epistemological survey in winter 2015. Table X shows the comparison between the postscores of the experimental group in fall 2014 (labeled in Table X as "pretest fall") as pretest and the post-test of the experimental group in winter 2015. There is an epistemological change in the CS factor towards

TABLE VII. A pairwise comparison of the variance of pre-post gains between the experimental group and the control group for the *source of knowledge authority* factor.

| Group (I) | Group (J) | Mean difference (I-J) | Standard error | Significant | 95% Confidence interval for difference | | Effect size |
|-----------|-----------|-----------------------|----------------|-------------|--|-------------|-------------|
| | | | | | Lower bound | Upper bound | |
| Cont | Exp | 0.088 | 0.076 | 0.248 | -0.062 | 0.237 | 0.11 |

TABLE VIII. The results of paired *t* test for the experimental group at the end of the winter semester 2015 (post-test) compared to the beginning of the fall semester 2014 (pretest).

| Factors | Mean pretest fall | Mean post-test winter | <i>N</i> | Paired differences ("Pretest"–"Post-test") | | Standard error mean | <i>t</i> | Significant (2-tailed) |
|-------------------------|-------------------|-----------------------|----------|--|--------------------|---------------------|----------|------------------------|
| | | | | Mean | Standard deviation | | | |
| Certainty or simplicity | 3.78 | 3.27 | 63 | 0.506 | 0.376 | 0.0474 | 10.7 | 0.000 |
| Justification: personal | 2.12 | 1.87 | 63 | 0.254 | 0.479 | 0.0598 | 4.24 | 0.000 |
| Source: authority | 3.87 | 3.38 | 63 | 0.488 | 0.556 | 0.0701 | 6.97 | 0.000 |
| Attainability of truth | 4.12 | 4.10 | 63 | 0.016 | 0.628 | 0.0792 | 0.20 | 0.842 |

TABLE IX. The results of paired *t* test for the control group at the end of the winter semester 2015 (post-test) compared to the beginning of the fall semester 2014 (pretest).

| Factors | Mean pretest fall | Mean post-test winter | N | Paired differences (“Pretest”–“Post-test”) | | | <i>t</i> | Significant (2-tailed) |
|-------------------------|----------------------|--------------------------|----|--|--------------------|------------------------|----------|------------------------|
| | | | | Mean | Standard deviation | Standard error mean | | |
| Certainty or simplicity | 3.59 | 3.59 | 52 | 0.000 | 0.354 | 0.0490 | 0.00 | 1.000 |
| Justification: personal | 1.90 | 1.73 | 52 | 0.173 | 0.552 | 0.0766 | 2.26 | 0.028 |
| Source: authority | 3.38 | 3.31 | 52 | 0.673 | 0.582 | 0.0807 | 0.834 | 0.408 |
| Attainability of truth | 4.00 | 3.81 | 51 | 0.186 | 0.591 | 0.0828 | 2.25 | 0.029 |

TABLE X. The results of paired *t* test for the experimental group at the end of the winter semester 2015 compared to the end of the fall semester 2014 (pretest).

| Factors | Mean pretest fall | Mean post-test winter | N | Paired differences (“Pretest”–“Post-test”) | | | <i>t</i> | Significant (2-tailed) |
|-------------------------|----------------------|--------------------------|----|--|--------------------|------------------------|----------|------------------------|
| | | | | Mean | Standard deviation | Standard error mean | | |
| Certainty or simplicity | 3.40 | 3.27 | 63 | 0.129 | 0.412 | 0.0518 | 2.49 | 0.016 |
| Justification: personal | 1.91 | 1.87 | 64 | 0.0391 | 0.394 | 0.0492 | 0.793 | 0.431 |
| Source: authority | 3.29 | 3.38 | 63 | -0.0952 | 0.521 | 0.0656 | -1.5 | 0.152 |
| Attainability of truth | 4.07 | 4.10 | 63 | -0.0317 | 0.581 | 0.0732 | -0.43 | 0.666 |

TABLE XI. The results of paired *t* test for the control group at the end of the winter semester 2015 compared to the end of the fall semester 2014 (pretest).

| Factors | Mean post-test fall | Mean post-test winter | N | Paired differences (“Pretest”–“Post-test”) | | | <i>t</i> | Significant (2-tailed) |
|-------------------------|------------------------|--------------------------|----|--|--------------------|------------------------|----------|------------------------|
| | | | | Mean | Standard deviation | Standard error mean | | |
| Certainty or simplicity | 3.59 | 3.59 | 52 | 0.000 | 0.354 | 0.0490 | 0.00 | 1.000 |
| Justification: personal | 1.90 | 1.73 | 52 | 0.173 | 0.552 | 0.0766 | 2.26 | 0.028 |
| Source: authority | 3.38 | 3.31 | 52 | 0.067 | 0.582 | 0.0807 | 0.834 | 0.408 |
| Attainability of truth | 4.00 | 3.81 | 51 | 0.186 | 0.591 | 0.0828 | 2.25 | 0.029 |

TABLE XII. A pairwise comparison between the experimental group and the control group (end of the winter 2015 semester) for the *certainty or simplicity of knowledge* factor in winter 2015.

| Pairwise comparisons | | | | | | | |
|----------------------------|-----------|-----------------------|----------------|-------------|--|-------------|-------------|
| Dependent variable: Change | | | | | | | |
| Group (I) | Group (J) | Mean difference (I-J) | Standard error | Significant | 95% Confidence interval for difference | | Effect size |
| | | | | | Lower bound | Upper bound | |
| Cont | Exp | -0.371* | 0.067 | 0.000 | -0.505 | -0.238 | 1.098 |

more expertlike behavior. There is no change in the JB, SK, and AT factors.

Table XI has the comparable scores for the control group. As you see there is no change in the CS and SK factors. There are some changes in the JB and AT factors towards less expertlike behavior. The experimental group and the control group for the CS factor are compared in Table XII. The difference is statistically significant, which means that students’ beliefs about certainty and simplicity of knowledge for the experimental group kept changing towards

more expertlike behavior in the winter semester and the combination of RW and Lbt influenced the students’ beliefs about certainty and simplicity of knowledge in two semesters. (A word of caution; the effect size found in this table may not be reliable because the control group and the experimental group at the end of the fall 2014 semester [initial values for the winter 2015 semester] had both changed and were no longer comparable.) From Table XIII, we see that the difference between the experimental group and the control group for the JB factor is

TABLE XIII. A pair wise comparison between the experimental group and the control group (end of the winter 2015 semester) for the *justification of knowledge* factor in winter 2015.

| Pairwise comparisons | | | | | | | |
|----------------------------|-----------|-----------------------|----------------|-------------|--|-------------|-------------|
| Dependent variable: Change | | | | | | | |
| Group (I) | Group (J) | Mean difference (I-J) | Standard error | Significant | 95% Confidence interval for difference | | Effect size |
| | | | | | Lower bound | Upper bound | |
| Cont | Exp | 0.337* | 0.119 | 0.005 | 0.103 | 0.572 | 0.43 |

TABLE XIV. A pairwise comparison between the experimental group and the control group (end of the winter 2015 semester) for the *attainability of truth* factor in winter 2015.

| Pairwise comparisons | | | | | | | |
|----------------------------|-----------|-----------------------|----------------|-------------|--|-------------|-------------|
| Dependent variable: Change | | | | | | | |
| Group (I) | Group (J) | Mean difference (I-J) | Standard error | Significant | 95% Confidence interval for difference | | Effect size |
| | | | | | Lower bound | Upper bound | |
| Cont | Exp | -0.035 | 0.123 | 0.774 | -.280 | 0.209 | 0.052 |

significant. The experimental group had stayed the same but the control group became less expertlike. We did not compare the SK factor because there was no change in either the experimental or control group. From Table XIV we see that the difference between the experimental group and the control group for the AT factor is not significant.

We also compared the results in the sections taught by the student author and those in the other sections. Table XV contains the results for the test of homogeneity of variance. We compare the variance of variance of pre-post gains. The high significance values are good because they mean we do have homogeneity of variance. We also ran an ANOVA test to see whether there is any difference between the experimental group taught by the student author and the experimental group taught by another lab instructor. In this case, because the significance value of 0.657 is greater than $\alpha = 0.05$, we accept the null hypothesis that says there is no significant difference between the groups. Finally, we conducted a multivariate test to confirm the results gained. We found that concerning the experimental groups for the CS factor there is no significant difference between group 1 and group 2. We had similar results for the control group for this factor and for the experimental and control groups in the JB factor. For the SK factor for the experimental groups

there was no significant difference between the groups. For this factor for the control groups there is a difference between the students taught by the student instructor and those taught by another instructor. There is more change in the epistemological beliefs of the students who were in the other lab instructor's sections and completed summary writing assignments. In the AT factor there was no significant difference between the groups.

4. Calgary study conclusions

The stronger statistics enabled by the larger data set generated by the Calgary study enabled a more in-depth and conclusive investigation of the effect of an exercise that had students reflect metacognitively on textual material before coming to class (RW) and then had interventions in class that had them examine subjects that produce cognitive dissonance. Specifically, the students' score on the most relevant CS metric clearly demonstrated that the experimental group made measurably more progress towards "expert-learning style" than the control group. In the Montreal or BC study the experimental group was assigned to work in C³G, and to write AE in addition to performing RW, whereas in the Calgary study, the amount of time on task was more equivalent because both groups performed

TABLE XV. The results for the test of homogeneity of variance between the groups taught by the student author and the groups taught by another lab instructor.

| Factors | <i>N</i> | Certainty or simplicity | Justification: personal | Source: authority | Attainability of truth |
|--|----------|-------------------------|-------------------------|-------------------|------------------------|
| Significance Value Experimental Groups | 110 | 0.157 | 0.907 | 0.669 | 0.122 |
| Significance Value Control Groups | 102 | 0.259 | 0.216 | 0.624 | 0.861 |

laboratorials. The difference between the two groups was that the experimental group performed RW and the control group performed summary writing. Thus, in this experiment it is clear that it is the addition of RW with an interactive intervention (Lbt) that was associated with a change in students' epistemological beliefs.

VI. OVERALL CONCLUSIONS

Our results from both the Montreal or BC study and the Calgary study provide a strong indication that a combination of an activity that gets students to examine textual material metacognitively with one or more interactive interventions can produce statistically significant epistemological change, in particular in the CS factor. Further, these gains are measurably stronger than those observed in the Control group. The breadth of this work indicates that such gains

have been observed across a range of institutions, student learning environments, and relevant pedagogical tools. We hope that our results would stimulate an effort by others to examine our hypothesis in other research settings.

There is a substantial body of research suggesting that interventions, which engage students in a process of reflection on their own beliefs, may assist in the development of epistemic beliefs [61–66]. Our study adds more evidence to this proposition. Hopefully, this study could inform the design of interventions for raising students' epistemic awareness in other domain areas.

ACKNOWLEDGMENTS

This work is supported by the Concordia University Open Access Author Fund.

-
- [1] R. R. Hake, Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses, *Am. J. Phys.* **66**, 64 (1998).
- [2] L. C. McDermott and Washington Physics Education Group, *Physics by Inquiry: An Introduction to the Physical Sciences* (John Wiley & Sons, New York, 1996).
- [3] K. Cummings and T. French, Development of a Problem Solving Assessment Tool for Introductory Physics Students, *Proceedings of the National Physics Education Research Conference* (Rochester, NY, 2001), pp. 87–91.
- [4] D. Hestenes, M. M. Wells, and G. Swackhamer, Force concept inventory, *Phys. Teach.* **30**, 141 (1992).
- [5] L. Ding, R. Chabay, B. Sherwood, and R. Beichner, Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment, *Phys. Rev. ST Phys. Educ. Res.* **2**, 010105 (2006).
- [6] L. Festinger, *A Theory of Cognitive Dissonance* (Stanford University Press, 1957); , Cognitive dissonance, *Sci. Am.* **207**, 93 (1962).
- [7] B. K. Hofer and P. R. Pintrich, The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning, *Rev. Educ. Res.* **67**, 88 (1997).
- [8] C. S. Kalman, Enhancing Students' Conceptual Understanding by Engaging Science Text with Reflective Writing as a Hermeneutical Circle, *Sci. Educ.* **20**, 159 (2011).
- [9] C. S. Kalman, S. Rohar, and D. Wells, Enhancing conceptual change using argumentative essays, *Am. J. Phys.* **72**, 715 (2004).
- [10] C. S. Kalman, S. Morris, C. Cottin, and R. Gordon, Promoting conceptual change using collaborative groups in quantitative gateway courses, *Am. J. Phys.* **67**, S45 (2009).
- [11] M. Milner-Bolotin, T. Antimirova, and C. S. Kalman, Comparison of the effectiveness of collaborative groups and peer instruction in a large introductory physics course for science majors, *Can. J. Phys.* **88**, 325 (2010).
- [12] D. Ahrensmeier, J. M. K. C. Donev, R. B. Hicks, A. A. Louro, L. Sangalli, R. B. Stafford, and R. I. Thompson, Laboratorials at the University of Calgary: In pursuit of effective small group instruction within large registration physics service courses, *Phys. Canada* **65**, 214 (2009).
- [13] T. L. McCaskey, *Comparing and contrasting different methods for probing student epistemology and epistemological development in introductory physics* (The University of Maryland Department of Physics, College Park, 2009).
- [14] A. Madsen, S. B. McKagan, and E. C. Sayre, How physics instruction impacts students' beliefs about learning physics, [arXiv:1403.6522v2](https://arxiv.org/abs/1403.6522v2).
- [15] E. F. Redish, J. M. Saul, and R. N. Steinberg, Student expectations in introductory physics, *Am. J. Phys.* **66**, 212 (1998).
- [16] W. Adams, K. Perkins, N. Podolefsky, M. Dubson, N. Finkelstein, and C. Wieman, New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey, *Phys. Rev. ST Phys. Educ. Res.* **2**, 010101-1 (2006).
- [17] C. S. Kalman, M. Milner-Bolotin, M. W. Aulls, E. S. Charles, G. U. Coban, B. [M.] Shore, T. Antimirova, J. Kaur Magon, X. Huang, A. Ibrahim, X. Wang, G. Lee, R. L. Coelho, D. D. N. Tan, and G. Fu, Understanding the nature of science and nonscientific modes of thinking in gateway science courses, in *Proceedings of the World Conference on Physics Education 2012, Ankara, Turkey*, edited by M. F. Taşar (Pegem Akademi, Ankara, 2014), ISBN:978-605-364-658-7, pp. 1291–1299.
- [18] B. K. Hofer and P. R. Pintrich, The Development of Epistemological Theories: Beliefs About Knowledge and Knowing and Their Relation to Learning, *Rev. Educ. Res.* **67**, 88 (1997).

- [19] L. B. Austin and B. M. Shore, Concept Mapping Of High And Average Achieving Students And Experts, *Eur. J. High Abil.* **4**, 180 (1993).
- [20] S. Pelletier and B. M. Shore, The gifted learner, the novice, and the expert: Sharpening emerging views of giftedness, in *Creative Intelligence: Toward Theoretic Integration* edited by D. C. Ambrose, L. Cohen, and A. J. Tannenbaum (Hampton Press, New York, 2003), pp. 237–281.
- [21] D. Tabatabai and B. M. Shore, How experts and novices search the Web, *Libr. Inf. Sci. Res.* **27**, 222 (2005).
- [22] *Standards for Foreign Language Learning, Preparing for the 21st Century* (American Council on the Teaching of Foreign Languages, Alexandria, 2001).
- [23] J. Discenna, *A Study of Knowledge Structure of Expert, Intermediate and Novice Subjects in the Domain of Physics* (Western Michigan State University Mallinson Institute for Science Education, Kalamazoo, 1998).
- [24] P. B. Kohl and N. D. Finkelstein, Expert and novice use of multiple representations during physics problem solving, in *Proceedings of the 2007 Physics Education Research Conference of the American Institute of Physics, Greensboro, August 2007* (American Institute of Physics, Melville, 2007), pp. 132–135.
- [25] J. D. Slotta and M. T. H. Chi, Overcoming robust misconceptions through ontological training, *Proceedings of the annual meeting of the American Educational Research Association, Montreal, Quebec, April 1999*.
- [26] D. Winter, P. Lemons, J. Bookman, and W. Hoese, Novice Instructors and Student-Centered Instruction: Identifying and Addressing Obstacles to Learning in the College Science, *J. Scholarship of Teaching Learning* **2**, 14 (2001).
- [27] W. G. Perry, Jr., *Forms of Intellectual and Ethical Development in the College Years: A Scheme* (Holt, Rinehart, and Winston, New York, 1970).
- [28] M. R. Belenky, B. M. Clinchy, N. R. Goldberger, and J. M. Tarule, *Women's Ways of Knowing: The Development of Self voice and Mind* (Basic Books, New York, 1986).
- [29] M. B. Baxter Magolda, *Knowing and Reasoning in College: Gender-Related Patterns in Students' Intellectual Development* (Jossey Bass, San Francisco, CA, 1992).
- [30] P. M. King and K. S. Kitchener, *Developing Reflective Judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults* (Jossey Bass, San Francisco, CA, 1994).
- [31] P. M. King and K. S. Kitchener, The reflective judgment model: Twenty years of research on epistemic cognition, in *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing*, edited by B. K. Hofer and P. R. Pintrich (Routledge, New York, 2002), pp. 37–62.
- [32] D. Kuhn, *The Skills of Argument* (Cambridge University Press, Cambridge, England, 1991).
- [33] D. Kuhn, R. Cheney, and M. Weinstock, *Cognitive Development* **15**, 309 (2000); P. F. Lazarsfeld, Interpretation of statistical relations as a research operation (R. Nice, Trans.) in *Social Research Methods: A Reader*, edited by C. Seale (Routledge, London, 2004), pp. 149–152.
- [34] M. Schommer, Effects of beliefs about the nature of knowledge on comprehension, *J. Educ. Psychol.* **82**, 498 (1990).
- [35] G. Qian and D. Alvermann, Role of epistemological beliefs and learned helplessness in secondary school students' learning science concepts from text., *J. Educ. Psychol.* **87**, 282 (1995).
- [36] D. Hammer and A. Elby, On the form of a personal epistemology, in *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing*, edited by B. K. Hofer and P. R. Pintrich (Routledge, New York, 2002).
- [37] B. K. Hofer, Dimensionality and Disciplinary Differences in Personal Epistemology, *Contemp. Educ. Psychol.* **25**, 378 (2000).
- [38] E. Borda, Applying Gadamer's concept of disposition to science and science education, *Sci. Educ.* **16**, 1027 (2007).
- [39] D. C. Finster, Developmental instruction: Part I. Perry's model of intellectual development, *J. Chem. Educ.* **66**, 659 (1989).
- [40] David C. Finster, Developmental instruction: Part II. Application of the Perry model to general chemistry, *J. Chem. Educ.* **68**, 752 (1991).
- [41] L. D. Bendixen, M. E. Dunkle, and G. Schraw, Epistemological Beliefs and Reflective Judgement, *Psychol. Rep.* **75**, 1595 (1994).
- [42] A. Elby, Helping physics students learn how to learn, *Am. J. Phys.: Phys. Educ. Res. Suppl.* **69**, S54 (2001).
- [43] M. U. Smith and L. C. Scharmann, Defining versus describing the nature of science: A pragmatic analysis for classroom teachers and science educators, *Sci. Educ.* **83**, 493 (1999).
- [44] S. A. Southerland, Epistemic Universalism and The Shortcomings of Curricular Multicultural Science Education, *Sci. Educ.* **9**, 289 (2000).
- [45] G. M. Sinatra, S. A. Southerland, F. McConaughy, and J. W. Demastes, Intentions and beliefs in students' understanding and acceptance of biological evolution, *J. Res. Sci. Teach.* **40**, 510 (2003).
- [46] C. L. Smith and L. Wenk, Relations among three aspects of first-year college students' epistemologies of science, *J. Res. Sci. Teach.* **43**, 747 (2006).
- [47] D. Hammer, Two approaches to learning physics, *Phys. Teach.* **27**, 664 (1989).
- [48] D. Hammer, Epistemological Beliefs in Introductory Physics, *Cognit. Instr.* **12**, 151 (1994).
- [49] D. Hammer and A. Elby, On the form of a personal epistemology, in *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing*, edited by B. Hofer and P. R. Pintrich (Erlbaum, Mahwah, NJ, 2002), pp. 169–190.
- [50] W. A. Sandoval and K. Morrison, High school students' ideas about theories and theory change after a biological inquiry unit, *J. Res. Sci. Teach.* **40**, 369 (2003).
- [51] D. Huffman and P. Heller, What does the force concept inventory actually measure?, *Phys. Teach.* **33**, 138 (1995).
- [52] P. Hewitt, Lessons from Lily on the Introductory Course, *Phys. Today* **48**, 85 (1995).
- [53] R. Havdala and G. Ashkenazi, Coordination of Theory and Evidence: Effect of Epistemological Theories on Students' Laboratory Practice, *J. Res. Sci. Teach.* **44**, 1134 (2007).

- [54] M. Schommer, Effects of beliefs about the nature of knowledge on comprehension, *Educ. Psychol. Rev.* **6**, 293 (1994).
- [55] D. J. Watters and J. J. Watters, Approaches to Learning by Students in the Biological Sciences: Implications for teaching, *Int. J. Sci. Educ.* **29**, 19 (2007).
- [56] H.-G. Gadamer, *Truth and Method* (Crossroads, New York, 1975/1960).
- [57] E. M. Nussbaum and L. D. Bendixen, Approaching and avoiding arguments: The role of epistemological beliefs, need for cognition, and extraverted personality traits, *Contemp. Educ. Psychol.* **28**, 573 (2003).
- [58] C. L. McDermott and P. S. Shaffer, *Tutorials in Introductory Physics* (Prentice Hall, Englewood Cliffs, NJ, 1998).
- [59] S. A. Radmacher and E. Latosi-Sawin, Summary Writing: A Tool to Improve Student Comprehension and Writing in Psychology, *Teach. Psychol.* **22**, 113 (1995).
- [60] B. D. Mowery, The Paired t-Test, *Pediatric nursing* **37**, 320 (2011).
- [61] J. Brownlee, N. Purdie, and G. Boulton-Lewis, Changing Epistemological Beliefs in Pre-service Teacher Education Students, *Teach. Higher Educ.* **6**, 247 (2001).
- [62] F. Cano, Epistemological beliefs and approaches to learning: Their change through secondary school and their influence on academic performance, *Br. J. Educ. Psychol.* **75**, 203 (2005).
- [63] N. Lyons, Dilemmas of Knowing: Ethical and Epistemological Dimensions of Teachers' Work and Development, *Harv. Educ. Rev.* **60**, 159 (1990).
- [64] S. Nist and J. Holschuh, Theory to Practice: Practical Applications of the Research on Epistemological Beliefs, *J. College Reading and Learning* **35**, 84 (2005).
- [65] P. Stacey, J. Brownlee, K. Thorpe, and D. Reeves, Measuring and Manipulating Epistemological Beliefs in Early Childhood Education Students, *J. Pedagogies and Learning* **1**, 6 (2005).
- [66] M. McLean, Can we Relate Conceptions of Learning to Student Academic Achievement?, *Teach. Higher Educ.* **6**, 399 (2001).