Replicating and understanding successful innovations: Implementing tutorials in introductory physics

N. D. Finkelstein and S. J. Pollock
Department of Physics, University of Colorado, Boulder, Colorado 80309, USA
(Received 17 May 2005; published 8 September 2005; corrected 20 September 2005)

We report a detailed study of the implementation of *Tutorials in Introductory Physics* at a large-scale research institution. Based on two successive semesters of evaluation, we observe students' improved conceptual mastery (force and motion concept evaluation median normalized gain 0.77, $N=336$), albeit with some student discontent. We replicate the results of original studies of tutorial effectiveness and document how and why these results occur. Additionally, using the Colorado Learning Attitudes about Science Survey we measure the support of students' expertlike beliefs about learning physics in our environment. We examine this implementation from a viewpoint that emphasizes varying contextual levels of this implementation, from students' engagement in individual tasks, to the situations in which these tasks are embedded, to the broader classroom, departmental, and educational structures. We document both obvious and subtle features that help ensure the successful implementation of these reforms.

DOI: 10.1103/PhysRevSTPER.1.010101 PACS number(s): 01.40.Fk

I. INTRODUCTION

Both popular[1] and research publications[2,3] document the increased performance of students who engage in research-based innovations in undergraduate physics courses, particularly the large-lecture introductory physics course. Predominantly, successes in these environments are reported by the authors of these reforms. Relatively recently, as these reforms have become more widespread, reports of replicating these original achievements in physics education reform are appearing.[4–8] These reports simultaneously document that it is possible for educators who are not the authors of reform to successfully implement them and identify some of the features that allow for or hinder the implementation of these efforts. The present work builds on this tradition of reporting the successful replication and adaptation of one of the most widely researched physics education research–(PER-)based reforms, *Tutorials In Introductory Physics*,[9] and applies a framework for understanding and organizing the features that shape when and why this implementation (and other reported replication studies) might be sustainable and successful.

Replication studies have been reported for some of the best-known reforms in the physics education research community, including studio physics, SCALE-UP, Workshop Physics,[4,6] Real-time Physics,[5] Interactive Lecture Demonstrations,[4] cooperative group problem solving,[4] and Peer Instruction.[7] Not only have these studies found that is it possible to achieve gains in conceptual mastery similar to those reported by the authors, but they also give some glimpses into why these reforms do and do not work. Such research makes clear that in studio physics reform, simply implementing the structure of small collaborative work and reduced lecture is not sufficient for improved student performance on conceptual surveys.[4,5] Research-based materials (in both subject content and pedagogical approach) are essential. Others document the importance of buy-in, or awareness of and belief in, the appropriateness of these new educational forms. Students arrive at these reform environments with different models of educational practice and often voice resistance to the educational changes and roles they are asked to serve.[4,5,7] Furthermore, Henderson identifies the instructor as critical to successful innovative practice.[10] Building on Rogers' model of dissemination of innovation,[11] Henderson notes that several significant factors shape one instructor's attempt to infuse reforms into his or her course: instructors awareness and understanding (including a robust instructional model and practical implementation strategies) and perceived external constraints.[10] At the same time, it is consistently recognized that "partial implementation" of complete educational reform leads to limited or no improvement in student conceptual mastery.[4,5,12] Each of these elements, content, pedagogical process, facilities, ease-of-use or fit with existing practices, and beliefs and buy-in from participants, is featured in successful implementations. Whether or not these are either necessary or sufficient elements for appropriate implementation of reform remains to be seen.

In the following case study of a secondary implementation of *Tutorials in Introductory Physics*, we triangulate a number of data sources to document when and how the local implementation of tutorials was successful. Applying a contextual constructivist model[13] we identify a framework for understanding which elements are critical, which must be adapted, and which must be reinvented.[14]

Tutorials in Introductory Physics[9] is a research-based curriculum, designed to supplement the conventional calculus-based introductory physics class by changing practices in the smaller recitation sections. The explicit goals are to develop student conceptual understanding and scientific reasoning skills. Many Ph.D. theses from students in the University of Washington (UW) Physics Education Group (PEG) have studied student learning difficulties and have been used to develop a comprehensive set of workbook exercises (along with associated pretests, homework sets, and follow-up exam questions). Each of these activities is designed to elicit student ideas and preconceptions, confront conflicts with nature, logic, and self-consistency, and then resolve these conflicts,

following a model of accommodation.[15,16] Students work in small groups, with the instructor(s) playing the role of "learning coach," asking guiding questions in a Socratic manner[3,17–19] without explicitly explaining the answers. Tutorials emphasize discussion and consensus as tools to resolve conceptual conflicts, focusing on qualitative understanding over traditional computational problem solving. A more complete description of the logistics, equipment, and instructor commitments is given below, and in the Tutorial Instructors Guide,[9] in Redish,[3] and at the UW PEG web page.[20]

The success of tutorials at the University of Washington is well documented to promote students' conceptual mastery of particular physics topics.[21–25] Primarily, these studies compare post-test results on exams (or interviews) with pretests on the same topic.[26] A frequently achieved informal goal of tutorials is that the undergraduate students should perform as well on post-test evaluations as the graduate physics teaching assistants (TAs) performed on pretest evaluations before going through the tutorial materials. We present a case study of a secondary implementation of tutorials at the University of Colorado (CU). Our study begins by documenting the logistics and effectiveness of this implementation, and then examines our implementation through a broader contextual lens, to better understand why and how these reforms work, and how to assess and improve their impact.

II. CU IMPLEMENTATION

A. Background and course structure

Tutorials were implemented first in the fall of 2003 in the introductory calculus-based physics course (Physics I) at the University of Colorado, Boulder (CU). CU is a large public research-extensive (primarily research-oriented) university.[27] The student population in the physics sequence described here is roughly 50% engineers and 20% natural science majors (20% are undeclared, and the remaining 10% are a broad mix of social science, arts, architecture, business, etc). The class is 75% male, and over half are freshmen. Three lectures per week are taught by the lead instructor (S.J.P.) in large halls which seat approximately 300 students. Multiple choice "concept tests" are used frequently in lectures (3–5 times/class) following Peer Instruction[12] methods, using a student personal response system.[28] Each week students attend an additional 50 minute recitation section staffed by a physics graduate student teaching assistant. All recitations in this physics course are scheduled on a single weekday, scattered among small classrooms with traditional seating for 28–32 students. This course serves 500–600 students each semester; the lectures are repeated twice, serving 250–300 students in the lecture hall at a time. There is no laboratory component associated with the first-semester course; however, students usually enroll in an independent laboratory course concurrent with their second semester of introductory physics. The content and syllabus of this course are largely defined by departmental tradition, and follow much of the first 15 chapters of Halliday, Resnick, and Walker[29] or a similar text during a 15-week semester. There are three evening midterm exams and a final exam, all equally weighted, which together account for 60% of the grade. In addition, conventional end-of-chapter problems are assigned and graded by a computerized system, LON-CAPA.[30] One of these homework questions is written up for hand grading by the TAs every week. Students are given the opportunity to work on homework collectively in the help room, a large TA-staffed room for all physics students to meet in and work on homework.

B. Structure of tutorials in introductory physics

Implementing tutorials involved the following components. One instructor spent two weeks visiting the University of Washington, participating in TA training classes and tutorials, and solicited advice from the UW Physics Education Group. Based on that advice, instructors (including the authors) constructed a small space with tables well suited to group work for half of the sections; the other half of the sections were still held in regular classrooms, with movable desks brought into small circles of four each. In spring 2004, the dedicated tutorial workspace was doubled, eliminating the need for the less desirable conventional classroom space. In order to adequately staff the tutorials, external funding[31] was secured to support an undergraduate learning assistant (LA) to couple with the departmentally funded graduate TA in each section, reaching a teacher:student ratio of 1:14 that approaches the suggested ratio of 1:10.[3] A weekly online tutorial pretest system[32] was set up. This pretest evaluation was required of all students every week[33] and functioned as an "elicitation" mechanism[15,16] for the students, while providing essential data on student preconceptions. Mandatory 75-minute preparatory meetings for the TAs and LAs were held each week to discuss the philosophy and practice of tutorials, including explicit guidance in the use of Socratic questioning as the primary tool for teacher-student interaction in the tutorials. Most of the preparatory sessions, however, were spent first examining student prior conceptions as represented on the pretests, and then working through the upcoming tutorial. The undergraduate LAs participated in an additional separate weekly one-credit-hour science education seminar through the School of Education.[34]

Each week during the 50-minute tutorial recitation sections, the Physics I students gathered in groups of four at tables, received brief instructions and graded work from the TAs, and then worked their way through tutorial exercises. Students were provided large scratch paper on the tables and appropriate equipment, and were strongly encouraged (by course instructor, TAs and LAs) to work through the tutorials collaboratively. Following the tutorial methodology, work during tutorials was never turned in and graded. Surveys and anecdotal evidence indicate that tutorials were often not completed in the 50 minutes allotted. Students were assigned 2 to 5 pages of homework each week from the tutorial[9] workbooks[35] which were graded by the TAs and constituted 15% of the overall course grade. Attendance at tutorial accounted for 4.5% of the final course grade. Twelve tutorials were run in each semester of implementation, largely following the published sequence.[36] Every midterm exam explicitly included a tutorial-related "long-answer question," which comprised 25% of the midterm exam. The course syllabus

and web page included basic explanations of the tutorial logistics, philosophy, and purpose. The implementation of tutorials was nearly identical for each of the two semesters included in this study. The only significant change was a doubling of physical workspace, so that no sections in the second term were in conventional classrooms with individual desks.

The details described in the paragraphs above represent the basic logistics for implementing tutorials within the broader institutional and departmental frameworks, training the instructor, construction of a physical space for the tutorials, providing a sufficient teacher-to-student ratio, training and preparing the TAs, assigning credits, and implicitly establishing a new set of expectations for students. In the second section of this paper we will use a contextual constructivist perspective[13] to examine more closely the role played by these aspects, and why and how tutorials fit into the course and impact student learning. First, we present data from the first two semesters of implementation (fall 2003 and spring 2004) to characterize the impact of this course structure on student learning and attitudes and beliefs about learning.

III. DATA AND RESULTS

Data were collected in two consecutive terms of the first semester of the introductory physics course and included measures of student conceptual gains, attitudes, and beliefs about physics and learning physics, and affective response to the tutorials and course.

A. Student conceptual learning

Conceptual learning in this course was extensively documented at a number of levels. Students completed weekly online pretests that included multiple-choice and free-response questions. Multiple-choice "concept tests" in lectures allowed the gathering of data on understanding of specific topics covered in tutorials. Questions were issued both in silent (individual) and talking (collaborative) formats, and were collected with a student personal response system.[28] Longer-term assessment of student conceptual mastery included analysis of student performance on tutorial homeworks and performance on exams. Well-documented questions from the University of Washington were used for most of the long-answer exam questions.[37] Additionally, multiple-choice questions designed to assess the impact of tutorials were added to each exam. Nationally normed conceptual surveys of student mastery were issued both pre- and post-instruction during the first two semesters of running tutorials: the force concept inventory[38] (FCI) in fall 2003 and the force and motion concept evaluation[39] (FMCE) in spring 2004. These were administered in the first and last recitation sessions of the term.

Student learning gains as measured by the FCI and FMCE tests fell into the "high-gain" category described by Hake.[40] Student performance on the FCI (given fall 2003) shifted from an average pretest score of 53% (median 50%) to an average post-test score of 81% (median 83%). The normalized gain of the averages, $g=(\text{post}-\text{pre})/(100-\text{pre})$, was 0.6. Matching individual students pre to post gives a median normalized gain of 0.67. The FMCE (given spring 2004) shifted from a pretest of 29% (median 23%) to a post-test of 74% (median 85%). The normalized gain of the averages $g=0.63$; the median gain matching students pre to post is $g=0.77$. Figure 1 shows the distribution of student performance for the spring 2004 pre and post FMCE. The most common score on the post-test was 100%.

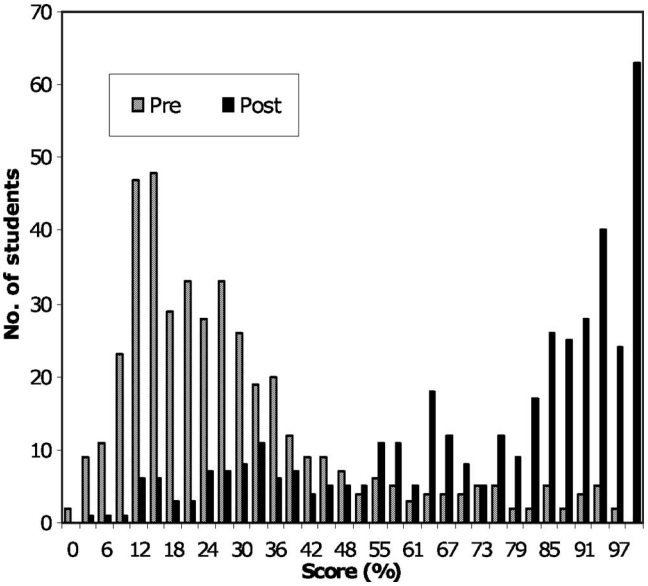

FIG. 1. Pre and post results for FMCE exam (spring 2004). Scoring of this exam follows Thornton (Ref. 41). All students who took the exam are shown in the histogram. Pretest ($N=486$) average 29%, median 23%. Post-test ($N=400$) average 74%, median 85%.

Because these conceptual assessments (the FCI and FMCE) examine only a very limited fraction of the content of our introductory physics course, we assessed student mastery of material covered on each tutorial (and in each week of class) with tutorial pretests and followup exam questions to provide more detailed measure of the impact of tutorials. Table I shows some representative student scores on measure of student mastery (in both the short term and longer term).

The data presented in Table I are a representative subset of the more complete assessment of the impact of the 12 tutorials. The "pre" column represents an average score for online multiple-choice questions given before the material was covered in tutorial. The data are averaged over several questions for each topic, and across the two implementations. Statistical errors in these data are typically a few percent, and semester-to-semester variation is also typically a few percent. Following the tradition of UW publications[21–25] we round all results to the nearest 5%. Participation in these pretests ranged from 75% to 90% each week for the more than 560 students who completed the course in fall 2003 and over 510 students in spring 2004. The "Post (click)" column lists averaged results of in-class concept questions answered with personal response systems the next class day following the tutorial. As may be observed from the table, the normalized learning gain on most of these topics is very high. In half of the topics shown in the table, the "Post (click)" score

TABLE I. Averaged scores (percent correct) on selected conceptual topics targeted by tutorials, rounded to the nearest 5%.

Topic	Pre	Post (click)	Post (exam)
Direction of **a** at top of ramp	20	70	75
Direction of **a** on curved path	10	70	65
Newton's third law	40	85	80
Galilean relativity:[a] acceleration	35	35	75
Sign of work	50	70	55
Compare $\|\Delta p\|$ (when objects stop vs recoil)	45	55	85
Use conservation of **p** to deduce **v**	20	80	70
F=m**a**, with extended body	20	45	70

[a]Galilean relativity was not covered in lecture or assigned as reading. The tutorial, the brief class discussion surrounding the in-class clicker question, and the associated tutorial homework were the sole treatments of this topic.

exceeds the average for the LA and TA pretest results for the tutorial in question. The shift from "Pre" to "Post (click)" is most attributable to the tutorials themselves as these are the only formal educational activity that students regularly engage in over this time period. Longer-term impact is seen by comparing the "Post (click)" (short-term) results to the exam (longer-term) results, given in the "Post (exam)" column. Once again, the "Post (exam)" column averages over several questions for each topic, and over all exams (both semesters, including finals) where the relevant questions appeared. These exam questions came weeks to months after the Tutorial.[42] Of course, there are many factors influencing student performance on these exams, including the tutorial and other homework problems, lecture, and the class discussion surrounding the question, all of which occur after the "Post (click)" question.

B. Replication

A primary concern when appropriating tested curricular reforms is the fidelity of implementation: how successful are we at replicating the original success of tutorials? One mechanism by which we may answer this is by a direct comparison of results with the original studies by University of Washington. On common long-answer questions issued on the midterm and final exams, we may compare answers. These results are shown in Table II. We compare the results of the UW implementation and the two semesters fall 2003 and spring 2004. The data show that we parallel published results from the primary implementation at UW.[21–25] In each of these cases we approximate the original results within 5% (the resolution of reported findings discussed above).[43]

C. Correlation studies

Many interventions were implemented in this course (online homework, a version of peer instruction with a personal response system, and tutorials), making it difficult to directly correlate student performance on measures of conceptual mastery with performance on (or attendance in) tutorials. For example, the correlation coefficient of FMCE normalized gain with tutorial homework is only 0.22, i.e., tutorial homework performance is associated with only a small fraction of the variance of the normalized FMCE gain. Nonetheless this positive correlation is statistically significant ($p<0.0001$). The modest value of the correlation coefficient may be due in large part to the clustering of many students with very high normalized gain (almost 30% of the class had normalized FMCE gains above 90%). Similarly, the correlation coefficient of tutorial attendance with FMCE gain is 0.12, again statistically significant but not large. Once again, this is not surprising, since average attendance (for students with measured FMCE gain) exceeds 95%. More telling is the correlation of tutorial homework score with overall course grade, shown in Fig. 2. Students who performed well on the tutorials (homework) did well in the course overall. Such data suggest that the tutorials are well integrated into the rest of the course, an idea that we return to below.

D. Students' attitudes and beliefs

In addition to the traditional content within any course, there are extensive sets of attitudes and beliefs about science that we teach to our students. The way we conduct our classes sends messages about how, why, and by whom science is learned. Such metamessages have been referred to as

TABLE II. Scores on selected conceptual topics measured on exam questions, rounded to the nearest 5%. (See Ref. 43.)

	UW (Ref. 25)		
	No tutorial	With tutorial	CU with tutorial
Atwood: tension before and after release	25	50	Spring: 55
Atwood: constrained motion, no friction	45	70	Spring: 45 (midterm) 75 (final)
Multiple horizontal objects: force diagrams	30	90	Spring: 95
Identify Newton's third partners	15	70	Fall: 80 / Spring: 70
Rank forces on different masses	60	60	Fall: 60
Of those who rank correctly: result of changing mass of middle object	35	>80	Fall: 80

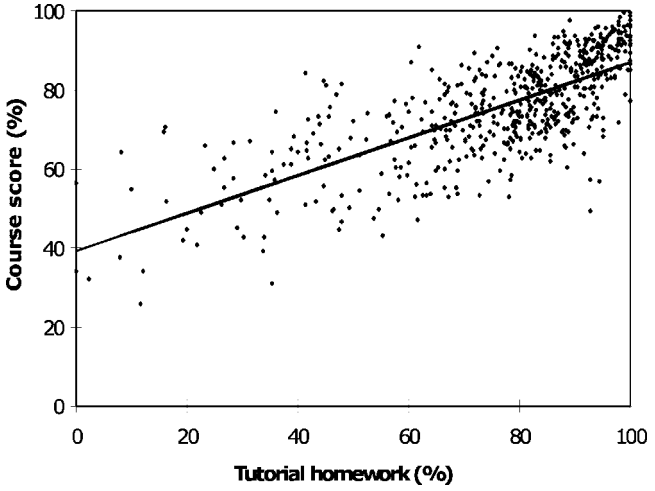

FIG. 2. Correlating overall performance in course (course score includes all graded elements except tutorial homeworks, expressed as a percent of total possible) with tutorial homework score. Data shown are for fall 03. The correlation coefficient is $r=0.73$ ($N=562$). The solid line shows a best fit (regression) to the data. Data for spring 2004 look very similar; the correlation coefficient is $r=0.65$ ($N=513$).

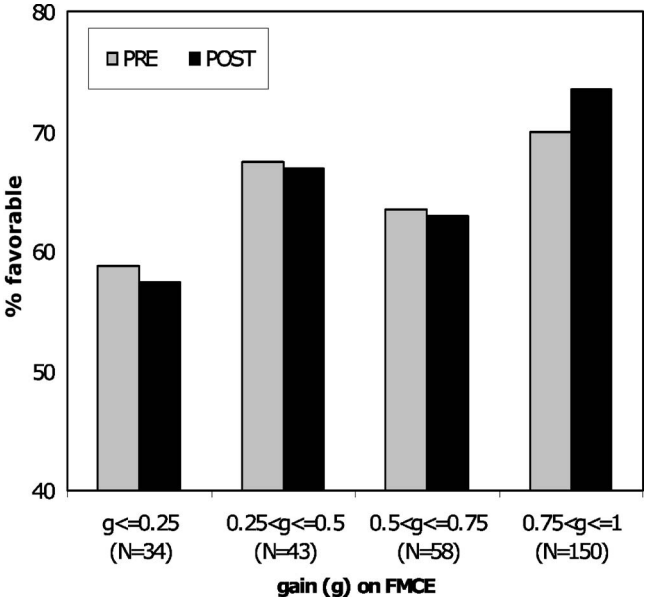

FIG. 3. Histogram of percent favorable responses in the CLASS "concepts" category pre and post for students with different normalized learning gains (g) on the FMCE.

the "hidden curriculum."[3] As offered in traditionally taught courses, some of the hidden curriculum is beneficial (e.g., the message that science is a coherent representation of the world) while other aspects are detrimental (e.g., the notion that women cannot be strong scientists). While decades of physics education research have reformed classroom practices to improve student mastery of conceptual domains,[2,3] these same class environments are often found not to improve student attitudes and beliefs (ABs). In fact, students are usually found to regress from more expertlike beliefs to more novice beliefs over the course of a semester.[3,45,45] In our findings and those of prior researchers, it is notable that such retreat in favorable student ABs is seen even for courses where reform pedagogy (including tutorials) is used and improved conceptual gains are observed. Although not intrinsically manifest in the curricular reforms we implement, there is potential for the development of this hidden curriculum, and it is this issue that we address next. To study the state and evolution of student beliefs in Physics I, we use a newly developed and validated instrument, the Colorado Learning about Science Survey (CLASS),[45] which is based in part on prior instruments used to evaluate student expectations.[44,46,47]

The overall result of the CLASS survey for each of the semesters of the reformed course showed no significant change in student beliefs from beginning to end of term–an improvement over the regress typically seen in introductory courses as measured by this and other instruments.[3,44,45] Because other implementations of Tutorials and tutorial-style approaches have documented the regress of student AB's,[3,44] it is likely that the lack of retreat we measure is associated with a more complex interaction of classroom instruction, and explicit attention to the hidden curriculum. In the second half of this paper, we will examine the role played by, and interactions among the various course elements, student beliefs, and practices surrounding the course. In Fig. 3, we show one example of the correlation between conceptual learning and the measured student attitudes about the role of concepts in learning physics. The histogram is binned based on normalized gain on the FMCE (low, moderate, high, very high) and within each bin, the pre and post scores (first and last weeks of the term) on a subset of CLASS questions that focus on the "Conceptual Understanding" category[48] are shown.[45] There are two interesting trends in the data. First, students with higher learning gains (as measured on the FMCE) tend to score more favorably (more "expertlike") on their attitudes and beliefs about the role of conceptual understanding in learning. Second, students in the lowest-gain group regressed to more novicelike beliefs over time, while students in the higher-gain groups improved. In the top bin (normalized learning gains on FMCE above 75%), the effect size of the pre-post CLASS shift is +0.2, a small effect. (For all students in the class, the effect size of the pre-post shift is negligible.) The Pearson's correlation coefficient of overall CLASS score (percent favorable) with final score on the FMCE is 0.24 (pre) and 0.34 (post), with nearly identical results using final course grades rather than FMCE scores. These correlation coefficients are statistically significant, indicating that attitudes and beliefs about learning may themselves influence, or be influenced by, the class structure and learning experience.

Reexamining these data by binning the overall score on the CLASS pretest in terms of percent favorable response (0="novicelike", 100="expertlike") provides another perspective. The highest gains on the FMCE are achieved by students with the highest CLASS prescores. This percentage of high FMCE gains declines as the student CLASS score decreases. A student is roughly five times more likely to achieve a very high FMCE gain (>0.8) if they begin the class with favorable expectations ($>80\%$) than with unfa-

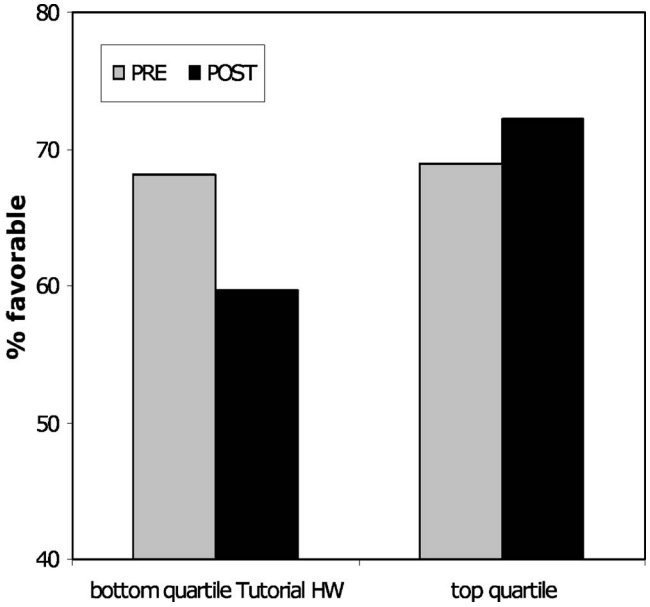

FIG. 4. Histogram of percent favorable responses in the CLASS "concepts" category pre and post for students in the bottom quartile for tutorial homework scores ($<72.8\%$) and in the top quartile ($>91.5\%$).

vorable expectations ($<40\%$). Similarly, students with unfavorable expectations about learning are roughly five times more likely to have low FMCE gains (<0.25) than those starting with favorable expectations. It appears that students who enter this course with more expertlike attitudes and beliefs have a better chance of making strong conceptual learning gains.[49]

In examining the relation between tutorial homework scores and students' CLASS beliefs (shown in Fig. 4), we observe similar trends as seen in Fig. 3, which compared CLASS performance as a function of FMCE gains. The lowest performers on tutorial homework make no gains (indeed, regress) in expertlike beliefs over the term, while the higher performers on homework improve somewhat in terms of expertlike beliefs. The difference in post-CLASS Concept score between the bottom- and top-quartile homework performers is large (>0.5 standard deviations), and statistically significant ($p<0.001$; two-tailed t test).

E. Student affect and evaluation

Students' affective responses (whether they enjoy the tutorials or not) are worth noting in a course that is their first experience in college science. Related to this sense of enjoyment is whether or not students believe the experience to be useful (whether or not, in fact, it is). Each week in class students were asked to report on various aspects of the course in online surveys. Participation in these extra-credit surveys averaged over 80% of the enrolled student body. Student responses to six relevant Likert-scale (agree or disagree) questions are listed in Table III.

Students have an overall favorable response to the utility of tutorials and the group work, as well as other course components and the course as a whole. Striking, however, is the sizable fraction of the class that does not enjoy the tutorial experience. At the same time students report enjoying the group work and talking about tutorials, suggesting the complex nature of affective response, students' reflection on past experiences, and how they identify the tutorials. (Traditional end-of-semester anonymous course and teacher evaluations were indistinguishable from prior semesters when the same instructor had not used tutorials.)

IV. DISCUSSION

A. Examining the how and why of implementing tutorials: A contextual constructivist analysis

While there are many lenses through which we may come to understand this secondary implementation of the tutorials at the University of Colorado, we explore the effects of the tutorials (what and why) from a contextual constructivist perspective.[50] Such an approach focuses on individuals in a social frame, rather than solely the individual and the cognitive processes associated with the individual. While cognition certainly depends upon individuals,[51] we argue that no less significant in the learning process are the environmental structures that constitute the educational experience. The lo-

TABLE III. Student responses to end-of-term online survey questions, rounded to nearest 5%. Questions were asked on a five-point Likert scale; "agree" in table corresponds to both "4" and "5." Totals do not sum to 100 because neutral "3" is a valid response.

Question (end-of-term survey)	Percent "agree"	Percent "disagree"
The tutorial is helpful for learning the material in this course	60	20
I enjoy tutorials	20	60
Working in groups and talking about physics in tutorial is helpful for learning the material	75	15
I enjoy working in groups and talking about physics in tutorial	70	15
The in-class ConcepTests are helpful for learning the material	95	0
I enjoy in-class ConcepTests	85	5

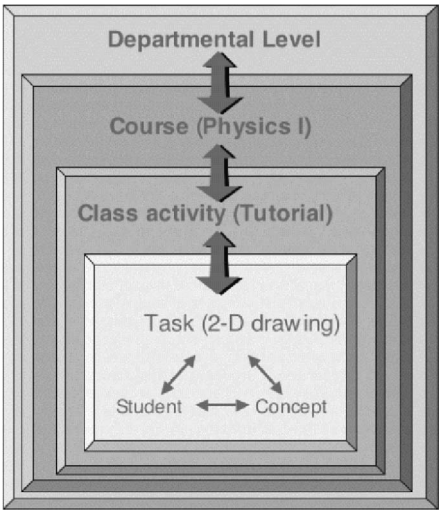

FIG. 5. Frames of Context: a particular task is embedded in a larger situation (a tutorial class activity), which is embedded in a larger microculture (a particular class), and so forth.

cal context, surrounding norms, and particular tools, whether a given tutorial lesson or 2D velocity arrows and a shared piece of paper used in that lesson, serve to mediate students' actions, and thereby participate in the intellectual processes we seek to promote.

A useful tool for identifying a particular level or scope of activity that shapes the success of tutorials is the notion of *frames of context*.[13] We might examine any number of frames: the tightest framing—a student engaged in a particular task, such as a student predicting the direction of an acceleration vector; a slightly broader frame—the situation surrounding a particular task, such as a given tutorial session on two-dimensional motion and acceleration; the local culture of the physics course or tutorial sessions, which is constituted by the collection of many such situations; the broader departmental and university norms; and so forth. Figure 5 illustrates the nested nature of these frames.

While these frames are useful for locating the level of detail we wish to examine, we also recognize that these levels influence one another. There are dynamic relations among constituent elements of the tutorial system, the associated physics course, and broader departmental, university, and environmental cultures. A tutorial session is not simply surrounding individual tasks, but constituted by them, as well as shaping the tasks themselves.[52]

B. Implementing tutorials: A frames of context perspective

In analyzing the various frames of context necessary for implementing tutorials, we highlight some of their key features:[13] Frames are nested and generally hierarchical (and outside more heavily influences inside than the reverse); they are dynamic, representing evolving relations among constituent elements; their boundaries are somewhat fluid: frames interact not only at their interfaces, but also via their constituent elements; and people participate simultaneously in many frames of context.

Below, the implementation of tutorials at Colorado is examined at each of four nested and increasingly broad frames: task formation, situation, course, and departmental frames.

1. Task formation

Because of the well-documented practices of the tutorials, it is relatively straightforward to replicate particular task formations. The particular tasks delineated by the tutorial workbooks and accompanying homework are designed simultaneously to be simple to implement and to be directed at improving students' mastery of foundational concepts in the domain. Because they are among the most studied educational reforms in college physics, the particular task formations have been demonstrated to be very effective at eliciting students existing ideas about the content domain, and at stepping students through a series of confronting or bridging activities to help the students build more expertlike understanding of the concepts.[16,53] The implementation of similar tasks at CU has led to similar results on assessments of students' mastery of particular concepts, as shown in Table II. These task formations, however, depend upon a variety of other constraints and interactions. Most notably, the students themselves are critical to the creation of effective task formations. While students' prior understanding of these domains has been thoroughly researched, and thereby included in the materials used in these tasks,[54] these studies have been conducted with a particular subset of U.S. students. Tutorials are designed for a calculus-based course and studies of student understanding have been conducted mainly at the University of Washington. Students who come from different backgrounds or contextual frames (cultural, linguistic, and gendered frames) bring different tool sets (norms, practices, beliefs, and understanding) for engaging in these tasks. In some cases, working with different student populations, it has been suggested that these tutorials are not as effective at promoting conceptual understanding as demonstrated by Washington or our current work.[8,55] One potential explanation for such findings is that the students themselves engage in, shape, and interpret these tasks in very different ways. Fortunately for the present study, the student demographic for students enrolled in calculus-based physics at University of Colorado is rather similar to the student population at the University of Washington. In particular, we observe similar pretest scores on common conceptual questions. Even so, in weekly observational notes collected by the learning assistants and through analysis of course surveys completed by students at the end of term, we have found that many students consider the tutorials an activity to be discharged rather than one that is to be learned from. Furthermore, as shown in Table III, students report not enjoying the tutorials. Such a view of these tasks likely comes from students' broader perspectives, which are shaped by the larger frames of context that include norms developed in other courses and the wider context of the education itself. Not surprisingly, tutorial tasks are strongly bound to these broader contexts, and most notably to situations—the conditions in which a particular task might take place.

Heuristics of implementing tutorials at the level of tasks. The particular tutorial tasks are well documented and easily

implemented (in terms of both infrastructure and practice) and hence exporting tutorials at this level is straightforward. Nonetheless, consistent with prior researchers, we claim the particular tasks are not sufficient to promote student learning; how these tasks are situated (whether or not students work alone, for example), and what goals are enacted in these tasks[3,47] (conceptual learning versus development of particular attitudes and beliefs) are both critical to successful implementation. Ensuring that these tasks are productive depends upon the broader situations in which they are embedded.

2. Level of situations

The environmental surrounding, the ensemble of task formations over time and the resources present for conducting a particular tutorial task, constitutes the surrounding situation. Critical elements for implementing productive situations include the design of a functional space, resources such as tables, papers (used as joint mediating tools), modest equipment, appropriate tests, and feedback,[56] and adequately trained staff. These features in collection allow for the students to productively engage in the tutorial tasks. However, at the same time, these elements themselves are critically dependent on the broader framings. The department has to commit to providing appropriate space and the added staffing requirements: roughly twice the usual number of assigned TAs is suggested.

From a pedagogical perspective, two forms of interaction are essential in these situations: student-student interaction and student-TA interaction. With the appropriate shift of resources, Colorado has doubled the number of rooms allocated to the recitation sections so that groups of students might be able to work at tables of four around tutorial materials and discuss the questions posted in the tutorial tasks. Students develop an expectation of talking as a form of educational practice (as opposed to passively sitting listening to a TA working through homework problems at a chalkboard in the front of the room).[57] It is worth noting that the tables do not force the students to interact, but do support the tasks that require interaction. In the first semester, when some of the students worked at individual desks, students were required to rearrange the environment physically (working against the original room design) to engage in these tasks, by bunching desks together. Had the desks been bolted down, the environment may well have precluded the forms of interaction we sought to support. Situations where students explain their reasoning to their peers, debate approaches, and collectively agree upon answers lead to enhanced student mastery of concepts (demonstrated above) and more sophisticated expectations about learning (as evidenced by modest increases, or no decreases in their beliefs). Notably, students also come to recognize the educational value of group work, and in particular such sentiments are captured in students' evaluation of the tutorials (Table III).

The student-TA interaction, which we advocate as essential in the productive implementation of tutorials, relies upon the existence of enough staff, and upon their adequate training and support. By having two leaders in a class of 24 students, the six groups of four students may each interact frequently enough with experts to ensure that their student-student interactions in given tasks are productive and support the broader course culture and norms. The tutorials build such interactions into the materials (such as the tasks of check-outs, where students check their answers with TAs before proceeding further in the tutorial) to support these productive learning situations. Meanwhile, adapting practices at the University of Washington, TAs and LAs must learn how to support and build these learning situations. The local tutorial leaders participate in relevant and intersecting frames of context in which they develop expectations and practices that support the tutorials. It is during a training session each week run by the course instructor that TAs and LAs develop the pedagogical content knowledge[58] appropriate for conducting tutorials. Furthermore, approximately half the tutorial leaders participate in an additional course on education and learning in physics or the sciences.[59] During these two forms of interaction, the tutorial leaders develop the norms, expectations, and an understanding of the practices in Tutorials. One of the critical norms of tutorials, emphasizing student responsibility and agency, is supported by TAs and LAs not disclosing answers during the tutorial sessions. Students are held responsible for understanding and explaining the content (on graded tutorial homeworks and targeted questions on exams). In the tutorial situations TAs and LAs emphasize these norms by supporting student reasoning and understanding more than verifying correct answers. As a result, students come to develop an expectation of learning and their roles in this process, as discussed in Sec. III D above. These participants and forms of interaction are critical in shaping the surrounding culture of tutorials.

Heuristics of implementing tutorials at the level of situations. How students engage in tutorial tasks (whether they simply discharge their duties of filling in answers or try to make sense of the concepts at hand) is significantly shaped by students' interaction with each other and with the TAs. Students support each other's understanding by making reasoning explicit; TAs support student engagement in tasks by *not* disclosing answers or grading student work in tutorial section (but do grade homeworks and exams). Other elements of the situation enable these productive interactions: locating students in the same room, providing appropriate desks for students to work at, eliciting student conceptions in advance of tutorials, and understanding that these ideas will be reconsidered on exams and homework. Each component of these situations is shaped by TA, student, and instructor understanding of the norms of this environment, which are (hopefully) based on a model of how students learn.

3. Course culture

Extending the field of view allows for examination of collections of many situations over time that are localized to a particular group of individuals. One way of describing this frame of context is by thinking of small-group cultures, or what Fine calls idiocultures.[60]

In many senses, developing a productive course idioculture is the goal of implementing tutorials at CU—we seek to create an environment which is the productive collection of particular norms, uses of tools (intellectual and material),

and their evolution over time. Researchers[61] have also referred to such as "habits of mind."[62] A wide variety of norms and practices are supported by the tutorials and surrounding course.

(1) Interactive, student-centered engagement: the tutorial sessions, peer instruction activities, and homework help sessions, where the students are responsible for identifying particular issues and answering questions they raise (often collectively).

(2) Students talking: related to a student-centered environment; however, tutorials and other elements (particularly the peer instruction and help room) shift students' beliefs about who is to be speaking during officially sanctioned class time.

(3) Not knowing, but discovering, is valued over answers: By not providing official solution sets, by not telling students of correct answers in section, but allowing students to discover and invent answers, the tutorials emphasize a more sophisticated epistemological frame.[63] This normative value is supported by not grading tutorials, and notably runs counter to much of traditional instruction and formal grading systems.

(4) The nature of science and scientific inquiry is valued explicitly along with particular content. Students are encouraged to focus on how we know what we know as much as they focus on what we know.

These norms are most clearly established by the course instructor and tutorial leaders. These norms are established through the practices, tasks, and situations described above and are embodied in many of the tools used in these practices. Sometimes these norms are stated overtly—as in the syllabus or in the grading rubrics—and sometimes they are left implicit in the particular tasks and situations in which students are placed. For example, the homework help room does not provide instructions on use, but engenders certain practices and norms such as student talk. While often directed by authority figures, these norms must be collectively established by students as well. If students do not accept or participate in these normative practices, the local classroom culture does not progress, no matter how insistent the instructor. When these norms sit at odds with existing norms and beliefs of students (developed in other educational environments) these idiocultures clash. For example, as observed by LAs and documented in surveys of student attitudes, to students who are committed to performing in lieu of understanding because they have been rewarded for such beliefs, the tutorials can seem a blur of busy work and students resist both the added work and nuanced approach of tutorials. They also report not enjoying tutorials. Carefully engineered educational environments, however, may foresee this and diffuse the distinction between performing and understanding by placing appropriate questions and assessments in the course (such as tutorial questions on the exams). As a result of such attention, we observe significant correlation between students' final grades (excluding their homework scores) and their performance on tutorial homework (Fig. 2). Nonetheless, such extrinsic focus remains a patch on the large problem of clashing ideology that is established from broader contextual frames such as the educational system's commitment to grading. We hypothesize that such a rift in expectations about learning and the practice of tutorials may be one of the causes of students' lack of enjoyment of tutorials. While the experience is evaluated as a positive learning environment (and this is borne out by evaluation of students' conceptual mastery), such learning experiences do not have to be enjoyable (like a potent medicine). Student affect is the subject of current and future research.

Without particular attention to these rifts in norms, students' expectations and beliefs about learning physics will tend to degrade. That is, students will come away from the course believing that physics has less to do with their lives outside the classroom than when they entered the course.[3,49] By interweaving the tutorials with the rest of the course (referencing tutorial activities in lecture, placing tutorial-style questions on exams, and linking these with homework), students at CU were encouraged to consider the coherent nature both of the course and of physics. Students hopefully recognize that tutorials do not sit as isolated or ancillary activity but as an integral activity to the course and to learning (or at least performing). The increase in student performance from the near-term ["Post (click)" questions, Table I] to longer-term evaluation (final performance; see Ref. 42) corroborates the integrated nature of tutorials with the course. Furthermore, those students who have a more expertlike understanding of conceptual framing and coherence of physics do better in the course, and tend to improve their understanding along these dimensions (see Figs. 3 and 4). At the same time, the content of the tutorials references activities previously covered, suggesting that physics is a coherent domain rather than a disconnected collection of trivia. While the tutorials themselves do not explicitly reference either the students' lives or particular ties between physics content and the world outside the classroom, both the instructor and the teaching and learning assistants regularly tied the tutorial concepts to real-world phenomena associated with students' lives. Finally, while subtle, the significant presence of and approach to student-TA interaction, e.g., not providing answers or solutions to tutorials, encouraged students to engage in and take charge of their learning, as well as introducing students to the nature of scientific investigation (supporting intellectual discovery over memorization).

Given these norms, one of the main challenges of defining and producing replication remains: how does one adopt and adapt a culture of norms and beliefs? The materials and even structural implementation of tutorials are relatively easy to replicate; however the norms surrounding these cannot simply be stated and exported—rather, they must be grown and adopted. That is, while it might be useful to explicitly state appropriate norms (as we attempt here), faculty must adopt these and appropriate them to make sense at their institutions. Just as students may learn in a lecture environment where information is stated, the learning process is one of construction of ideas and understanding, so too do faculty need to build an understanding of supportive norms of tutorials at their institution even if the norms are overtly stated. As a result, UW emphasizes the importance of visits to Washington for workshops or training sessions at national conferences. Just as LAs develop an ideological frame and norms by participating as they do, so too might faculty. Because the lead instructor had visited the University of Washington and because other researchers in the CU community

had spent time with the tutorial authors and studied their approach, adapting an appropriate ideological frame was more likely in the CU implementation. Such an adoption of ideology critically depends upon the broader framing of faculty worlds, of course.[64]

Heuristics of implementing tutorials at the level of course culture. The norms and expectations established by and for the participants in the tutorials are among the most critical and subtle features of implementing the reforms. The situational interactions described in the previous section are guided by this localized culture. The tutorials are an integrated set of activities, including pretests, tutorial tasks, TA training, homework sets, and exams,[9] all of which establish a set of beliefs about learning physics in this course. It is through the norms established in this system of activities accompanied by other explicit efforts in lecture that students and TAs learn critical features of productively engaging in the tutorials. For instance, ideas are constructed by the students, and answers are reasoned and justified as opposed to corrected. It is also possible in these environments, though by no means inevitable, that productive beliefs about learning physics are established. For example, tutorials and physics are not disconnected from the rest of class or the world at large. Because the norms and expectations evolve over time, adopting a model of enculturation[59] of participants (from instructor, to TAs, to students) appears to be effective for establishing a supportive culture.

4. The broader lens: Department, university, and social frame

While various broad frames shape the tutorial experience, including department, university, and the larger national educational system, these are collapsed for the sake of brevity. A number of competing interests and norms arise in these broader educational frames: cost of implementation, teaching and learning culture, and what each of these means. Twice the number of TAs is required to run the tutorials as is assigned by the department. While addressed in different ways at UW and CU, increased teacher:student ratio is required for appropriate tutorial experiences. At CU, through external funding (an NSF grant) and a novel course on teaching and learning physics,[65] the tutorials are adequately staffed and prepared. Fundamentally, however, the existence and training of these tutorial leaders (alternative to the traditional TA models) requires support from the departmental or broader institutional levels to remain sustainable. Once (if) external funding is withdrawn (yet a broader framing on local practice), alternative models need to be in place. Two such alternatives are commitments from the department for additional funding per student, or expanding the class on teaching and learning physics. Each provides the added lead support required, one through funding, and the other through a service-learning educational experience.[66] Meanwhile, other resources, equipment, faculty time, and space require some funding and institutional commitments; however, these are more closely aligned with the levels of commitment to education that are currently provided.

More subtle, but no less significant, is that what counts for education shapes the broader commitment to tutorials. In traditional institutions, teaching evaluation is at best loosely coupled with student learning. Most institutions, including CU, utilize the end-of-course questionnaire as the predominant tool for evaluating faculty teaching practice. Plenty of discussion and some research has examined the limitations of such tools.[67] However, it is generally recognized that these course questionnaires often do not reflect measures of student learning. Simultaneously, students themselves have been embedded in an educational system valuing performance (distinct from learning) and have come to value evaluation more than understanding. Thus, students have fairly well-established beliefs about what practices constitute an appropriate educational experience. When new and different forms of practice are introduced, especially those that require more effort, students often resist. These forms of resistance may be realized in each frame of context discussed above, from task (dis)engagement to no longer participating in the course, or even providing negative feedback on the course questionnaire, which in turn shapes faculty practice. Thus, while faculty beliefs and goals may be aligned with reformed practices, systemic barriers to implementing these practices may prevent such efforts.[64]

Optimistically, tutorials and related PER-informed activities can affect the broader institutional norms by developing sets of expectations among the participants in the educational system. Students and faculty may begin to reshape their expectations of their roles in a class, and tools such as surveys of students' mastery of concepts or beliefs may begin to become staple practices in the departmental use. Seeds of such a transition may be observed at the University of Colorado where faculty are beginning to more broadly adopt pre- and post-tests and discuss on a regular basis what the surrounding goals and measures of success are for the suite of undergraduate courses.

Heuristics of implementing tutorials at the level of department and institution. Tutorials require the support of the department or broader institution. The traditional model of course support does not provide the necessary resources required for sustainable implementation of tutorials, which requires more space, staff, and training. There are many mechanisms by which such support might be provided, such as external or internal grants, TA training courses, and/or leveraging technology and human resources. These added resources allow the tutorials to operate as originally designed. Simultaneously, effective PER reforms can feed back to the broader system and benefit not only individual students, but also departmental and institutional efforts in education, such as revising forms of teaching evaluation. Such positive feedback potentially justifies the added departmental and institutional commitments.

V. CONCLUDING REMARKS

Clearly, it is possible to replicate the success of tutorials demonstrated by the curricular authors. Notably, however, such replication is more than a matter of repeating particular activities or worksheets. It is a matter of appropriating activities and surrounding structures: the particular tasks, surrounding situations in which tasks are embedded, and the larger norms, practices, and infrastructure required to support

the tutorial situations. Because these activities must fit into new environments (that is, the boundaries at each of these level of detail must match boundary conditions), there is a mix of adoption, adaptation, and reinvention.

In addition to documenting the processes of implementing tutorials, we apply a different framework for understanding this educational practice in context, referred to as "contextual constructivism." This framework specifically attends to the social and contextual nature of educational environments and student learning. The idea of *frames of context* provides a mechanism by which we might understand how nested features of these environments (from departmental norms to specific learning tasks) shape and are shaped by the participants. We exploit this framing to identify particular levels of context that are relevant—e.g., the need for tables to support certain tasks, such as student discussion of a given concept—and how these levels are co-constituting—e.g. student talk supports a norm of tutorials: that students are the agents of their own learning and that reasoning is as important as final answers. Furthermore, we identify how particular frames of context support or sit at odds with preexisting frames, such as those that encourage students to focus on answers more than reasoning. In short, these frames of context provide an organizing mechanism for thinking about which features of the system are critical for success, as measured at whichever level of interest—from that of students (are they learning?) to that of the broader institution (is such a model sustainable?). In applying this model we simultaneously emphasize the critical role of context in student learning and provide a mechanism to build and understand environments supportive of learning.

While we have demonstrated some success at improving students' conceptual mastery and developing productive attitudes and beliefs about physics as a result of implementing tutorials, several challenges remain. First, we seek to further develop our implementation of these activities so that students' affective responses are more reflective of the joy and productivity of these intellectual tools and processes that we as scientists and educators experience. Next, further investigation is needed to address the increasingly diverse student populations we seek to serve in our school system. While rare, there are some reported investigations into the application and applicability of tutorials and other PER-based reforms for underrepresented populations.[8,55] Finally, while we demonstrated that tutorials are scalable (or replicable) a major question of sustainability remains. Will other faculty adopt these practices? Do students develop a more positive response to these practices over time? Are there models of sustainable support for the tutorials and the associated increased demands on the educational institution?

ACKNOWLEDGMENTS

We gratefully acknowledge the support of Lillian McDermott, Peter Shaffer, Mackenzie Stetzer, and the University of Washington Physics Education Group; Kathy Perkins, Wendy Adams, Patrick Kohl, Michael Dubson, and the Physics Education Research Group (PER@C) at Colorado. This work has been supported by the National Science Foundation (Grant No. DUE-CCLI 0410744), the American Physical Society (Colorado PhysTEC program), and the University of Colorado.

[1] R. Panek, 101 redefined. *New York Times, Educational Life Supplement*, January 16, 2005.
[2] L. C. McDermott and E. F. Redish, Am. J. Phys. **67**, 755 (1999).
[3] E. F. Redish, *Teaching Physics with The Physics Suite* (Wiley, New York, 2003).
[4] K. Cummings, J. Marx, R. Thornton, and D. Kuhl, Am. J. Phys. **67**, S38 (1999).
[5] M. C. Wittmann, http://perlnet.umephy.maine.edu/research/mcw.htm
[6] J. W. Belcher, MIT Faculty Newsletter **16**, 8, (2003).
[7] A. P. Fagen, C. H. Crouch, and E. Mazur, Phys. Teach. **40**, 206 (2002).
[8] M. S. Sabella, in Proceedings of the 2002 Physics Education Research Conference edited by S. Franklin, K. Cummings, and J. Marx (PERC Publishing, New York, 2002), p. 79, piggy.rit.edu/franklin/PERC2002/Sabella.pdf
[9] L. C. McDermott and P. S. Schaffer, *Tutorials in Introductory Physics* (Prentice-Hall, Upper Saddle River, NJ, 2002).
[10] C. Henderson, Am. J. Phys. **73**, 778 (2005).
[11] E. M. Rogers, *Diffusion of Innovations* (Free Press, New York, 2003).
[12] E. Mazur, *Peer Instruction: A User's Manual* (Prentice-Hall, Upper Saddle River, NJ, 1996).
[13] N. D. Finkelstein, Int. J. Sci. Educ. (to be published), physics/0505092 (retrieved July 4, 2005).
[14] The ideas of adoption, adaptation, reinvention, and invention are borrowed from Henderson. (Ref. 10).
[15] G. J. Posner, K. A. Strike, P. W. Hewson, and W. A. Gertzog, Sci. Educ. **66**, 211 (1982).
[16] L. C. McDermott, Am. J. Phys. **61**, 295 (1993).
[17] R. R. Hake, Phys. Teach. **30**, 546 (1992).
[18] R. A. Morse, Phys. Teach. **32**, 276 (1994).
[19] A. Collins, in Schooling and Acquisition of Knowledge, edited by R. C. Anderson, R. J. Spiro, and W. E. Montague (Lawrence Erlbaum, Hillsdale, NJ 1977).
[20] http://www.phys.washington.edu/groups/peg
[21] D. E. Trowbridge and L. C. McDermott, Am. J. Phys. **49**, 242 (1981).
[22] D. E. Trowbridge and L. C. McDermott, Am. J. Phys. **48**, 1020 (1980).
[23] R. A. Lawson and L. C. McDermott, Am. J. Phys. **55**, 811 (1987).
[24] L. McDermott and P. S. Shaffer, Am. J. Phys. **60**, 994 (1992).
[25] L. McDermott, P. S. Shaffer, and M. Somers, Am. J. Phys. **62**, 46 (1994).
[26] Researchers at the University of Maryland have found similar results. Using the UW tutorial approach described above, Redish and colleagues demonstrated that with modifications involving microcomputer-based laboratories, students demonstrated significantly greater improvement on the Force Concept Inventory

than those students who did not use Maryland tutorials.

[27] http://www.carnegiefoundation.org/classification/cihe2000/defnotes/definitions.htm

[28] http://www.h-itt.com/

[29] D. Halliday, R. Resnick, and J. Walker, *Fundamentals of Physics*, 6th ed. (John Wiley and Sons, New York, 2001).

[30] http://www.lon-capa.org

[31] STEM Colorado is part of the NSF Science Technology Engineering and Mathematics (STEM)-Teacher Preparation program.

[32] These pretests were served over the web and hosted by the University of Washington Physics Education Group. They are "pre" tutorial (but generally after clasroom or lecture instruction).

[33] Student participation on pretests accounted for 2.5% of their overall course grade.

[34] As part of the NSF-funded STEM-TP program, LAs enroll in a course, Educational Philosophy, Theory, and Practice, offered by the School of Education.

[35] These build on ideas and concepts raised in the in-recitation tutorial activity.

[36] Specifically, weekly tutorials covered acceleration in one dimension (1D), motion in two dimensions, forces, Newton's second and third laws, tension, work and the work-energy theorem, relative motion, a modified conservation of momentum in 1D (modified version provided by UW PEG), rotational motion, dynamics of rigid bodies, simple harmonic motion (modified version provided by UW PEG), and superposition and reflection of pulses.

[37] These questions came both from published work (Refs. 21–25) and from their large database of questions used to develop and assess the impact of tutorials.

[38] D. Hestenes, M. Wells, and G. Swackhamer, Phys. Teach. **30**, 141 (1992).

[39] R. K. Thornton and D. R. Sokoloff, Am. J. Phys. **66**, 338 (1998).

[40] R. R. Hake, Am. J. Phys. **66**, 64 (1998).

[41] R. K. Thornton, D. Kuhl, K. Cumming, and J. Marx (private communication).

[42] Notably, most in-class ("click") questions are answered collaboratively. Asking a question individually, hiding the results, and immediately following up with a 2-minute conversation typically increases scores by about 20–40 % (Ref. 12). The exam results, which are answered individually, therefore represent a larger gain than may first appear.

[43] In the UW data, the authors present the percentage of the class reporting the correct answer with the correct reasoning. The CU data report the average score of students (graded for correctness including reasoning). While these differing approaches could lead to different findings for differing statistical distribution of student responses, we have observed on sample questions that our measure of student scores generally matches the measure of the fraction of students scoring correctly on these questions, within the reported measurement uncertainty of 5%.

[44] E. F. Redish, J. M. Saul, and R. N. Steinberg, Am. J. Phys. **66**, 212 (1998).

[45] W. Adams, K. Perkins, N. Finkelstein, and C. Wieman, Proceedings of the 2004 Physics Education Research Conference, edited by J. Marx, P. Heron, and S. Franklin, AIP Conf. Proc. (to be published).

[46] I. Halloun, in *The Changing Role of Physics Departments in Modern Universities*, edited by E. F. Redish and J. S. Rigden, AIP Conf. Proc. No. 399 (AIP, Woodbury, NY, 1997).

[47] http://www2.physics.umd.edu/elby/ebaps/home.htm

[48] The concepts category refers to a portion of sense-making activity where understanding physics includes understanding the broader conceptual framing of the problems at hand. Questions include, "When I solve a physics problem, I explicitly think about which physics ideas apply to the problem" (Ref. 45).

[49] K. K. Perkins, W. K. Adams, S. J. Pollock, N. D. Finkelstein, and C. E. Wieman, in Proceedings of the 2004 Physics Education Research Conference edited by J. Marx, P. Heron, and S. Franklin, AIP Conf. Proc. (to be published).

[50] Such a perspective draws from the sociocultural school of psychology and learning and particularly draws from the work of Cole [M. Cole, *Cultural Psychology: A Once and Future Discipline* (Harvard University Press, Cambridge, MA, 2001)].

[51] A variety of perspectives on individual cognition, from a resources model [A. A. di Sessa, Cogn. Instruct. **10**, 105 (1993)] to classical mental constructs [D. E. Ruumethart, in *Theoretical Issues in Preceding Comprehension*, edited by R. J. Spono, B. C. Bruce, and W. C. Brewer (Erlbaum, Hillsdale, NJ, 1980)], have been quite productive in the PER community for developing models of student learning in physics.

[52] Furthermore, it is worthy of note that students participating in tutorials participate in many other contextual frames. Participants, students, TAs, LAs, and the instructor are all members of various cultural groups, family structures, and social circles, each with it own sets of norms, beliefs, situations, and particular tasks. One of our goals in the course is to recognize and value these historically developed and productive sets of norms and practices [L. Moll, C. Amanti, D. Neff, and N. Gonzalez, Theory Pract. **31**, 132 (1992)]. While we could analyze any number of sets of nested frames of context, in order to analyze critical elements of the tutorials, we find it useful to consider a particular perspective common to the students, instructor, and broader institution—notably the academic frame.

[53] J. Clement, J. Res. Sci. Teach. **30**, 1241 (1993).

[54] That is, not only do the material artifacts of the tutorial worksheets carry with them a particular representation of physics (well vetted by science and history) and pedagogy, but carry with them an understanding of students (what they know and how they know). For example, the particular tasks are designed to elicit common conceptions of the world that students have. A second example is that there are few references to real-world or personal activities. This approach promotes a more abstracted and disconnected way of knowing, which addresses different populations of students differently [M. Dancy, in Proceedings of the 2003 Physics Education Research Conference, edited by J. Marx, S. Franklin, and K. Cummings, AIP Conf. Proc. No. 720 (AIP, New York), p. 39].

[55] F. J. James, in Proceedings of the 129th AAPT National Meeting, Sacramento, CA, 2004 (unpublished).

[56] It is at the level of situation that we might understand why we might only compare UW and CU student conceptual mastery to 5%. While particular questions might be identical, the situation in which these questions are asked [where these questions are asked (exam, or in class), the question order (which questions precede and follow), and answer order within the questions] all affect student performance (Ref. 3). While the particular

tasks may be similar (answer a particular multiple-choice question), the situations in which these occur vary dramatically enough to recontextualize the tasks (and hence student performance).

[57] Such expectations arise from and shape the surrounding culture of the tutorials and physics course. Notably, while productive for situations conducive of educational tasks, such expectations sit at odds with other educational cultures in which students have participated.

[58] L. S. Shulman, Harv. Educ. Rev. **57**, 1 (1987).

[59] Notably the inclusion and training of the tutorial leaders occurs through an apprenticeship model [J. Lave and F. Wenger, *Situated Learning: Legitimate Peripheral Participation* (Cambridge University Press, New York, 1991)], whereby they begin to take-ownership of the activities that occur.

[60] G. A. Fine, *With the Boys* (University of Chicago Press, Chicago, 1987).

[61] P. Hutchings and L. S. Shulman, Change **31**, 10 (1999).

[62] Though these "habits" might be considered more global than simply existing within a microculture such as a course.

[63] D. Hammer, A. Elby, R. E. Scherr, and E. F. Redish, in *Transfer of Learning: Research and Perspectives*, edited by J. Mestre (Information Age Publishing, Greenwich, CT, 2005).

[64] M. Dancy and C. Henderson, Proceedings of the 2004 Physics Education Research Conference edited by J. Marx, P. Heron, and S. Franklin, AIP Conf. Proc. (to be published).

[65] N. D. Finkelstein, J. Coll. Sci. Teach. **33**, 37 (2003).

[66] N. D. Finkelstein, Jou. Scholar. Tech. Lear. **4**, 1 (2004).

[67] R. R. Hake, physlrnr archives http://listserv.boisestate.edu/archives/physlrnr.html, May 14, 2004, and references therein.