College physics students' epistemological self-reflection and its relationship to conceptual learning

David B. May
Department of Physics, The Ohio State University, Columbus, Ohio 43210

Eugenia Etkina
Graduate School of Education, Rutgers, The State University of New Jersey, New Brunswick, New Jersey 08901-1183

(Received 22 April 2002; accepted 5 July 2002)

Students should develop self-reflection skills and appropriate views about knowledge and learning, both for their own sake and because these skills and views may be related to improvements in conceptual understanding. We explored the latter issue in the context of an introductory physics course for first-year engineering honors students. As part of the course, students submitted weekly reports, in which they reflected on how they learned specific physics content. The reports by 12 students were analyzed for the quality of reflection and some of the epistemological beliefs they exhibited. Students' conceptual learning gains were measured with standard survey instruments. We found that students with high conceptual gains tend to show reflection on learning that is more articulate and epistemologically sophisticated than students with lower conceptual gains. Some implications for instruction are suggested.
[DOI: 10.1119/1.1503377]

I. INTRODUCTION

There is growing recognition in research about student learning that students' epistemologies play an important role in helping them construct knowledge. The word "epistemology" is used in many different ways.[1] When applied to student learning, we can understand student epistemologies as their beliefs or views about how knowledge is constructed and evaluated. These beliefs have always received attention from researchers, and interest among both teachers and researchers has increased dramatically in recent years.[2]

The use of interactive curricula and pedagogical techniques based on research in physics education is also on the rise, especially in the introductory courses. Many instructors and researchers have focused on developing and assessing student conceptual understanding and problem-solving skills in these courses. At the same time, little has been done to measure or address students' epistemological beliefs in these settings, or their possible relationship to conceptual and problem-solving knowledge.

Epistemological beliefs may depend on the specific content domain. If so, measurements in a variety of conceptual areas will have a better chance of more accurately reflecting the diversity and consistency of the beliefs. Beliefs might also depend on the particular way in which they are measured. Questionnaires and interviews are commonly used to study students' epistemologies, but neither is an integral part of the instruction; other measurements may provide results that are more relevant to instruction.

This study makes use of Weekly Reports,[3] open-ended journals that have been used in a number of different courses. In these reports, students reflect on what they learned in a given week and how they learned it, by responding to specific questions. Weekly Reports provide a context for research into students' epistemological beliefs about different areas of physics knowledge. The reports were used for two quarters with about 200 students in an existing introductory physics course for honors engineering majors.

By analyzing how students reflect on their learning in these weekly journals and by using standard measures of conceptual understanding and problem-solving ability, we hope to begin to answer three questions related to epistemological beliefs and self-reflection:

(1) What do students prefer to describe while reflecting on their learning? (We call these "epistemological preferences.")
(2) How might the amount of reflection they exhibit relate to their conceptual learning gains?
(3) How might their epistemological preferences relate to their conceptual learning gains?

By examining these relationships, we hope to begin to understand the kinds of reflection that are most productive for conceptual learning in physics and to explore the implications for instruction that they suggest.

The theoretical underpinnings of the study are explored in Sec. II. Section III explains features of the study's design, including our sample selection method and our coding scheme for analyzing the Weekly Reports. In Sec. IV we present our qualitative and quantitative findings. A summary and conclusions are presented in Sec. V, and in Sec. VI we discuss the implications of the results for instruction.

II. REVIEW OF THE LITERATURE

Researchers have characterized personal epistemological beliefs in several different ways. Most include beliefs about how knowledge is constructed and evaluated and how knowing occurs.[4] Most have chosen to think of epistemology as a sequence of developmental stages[5] or as a few orthogonal dimensions.[6,7]

Studying epistemological beliefs is important because they influence motivation[8] and affect the selection of learning strategies by students.[7] In particular, immature beliefs affect students' ability to integrate their understanding of science concepts.[9] Epistemology may also affect the ways that students evaluate their learning.[10,11]

Most studies have assumed the domain and context independence of epistemological beliefs.[8] As Hammer and Elby[12] have pointed out, however, research indicates that beliefs depend on the content *domain* in question[13] (such as discipline or topic within a discipline) and on the particular *context* of the belief[8,14] (such as a class discussion, solo problem solving, or a written survey). Different resources for learning may be cued in different domains and contexts. Hofer and Pintrich[4,8] also call for more domain-specific research into epistemological beliefs, and for it to be situated in more naturalistic contexts. This study aims to fill this need for epistemology research within more specific physics domains and in a particular context, that of regular self-reporting of learning.

Thus far, only a few researchers have explored student epistemologies in the domain of introductory physics.[15–17] Hammer's framework,[16] the result of extensive interviews, includes dimensions for beliefs about the structure of physics knowledge, the content of physics knowledge (formulas versus concepts), and the process of learning physics. These inter-related dimensions are part of the basis for the Maryland Physics Expectations (MPEX) survey, a multiple-choice instrument designed for introductory physics classes.[17] Results of using the MPEX in a wide variety of physics classes suggest that most students, unlike physics experts, maintain beliefs that knowledge consists of disconnected facts and formulas, handed down by authority, that are unrelated to their everyday experiences.

These results confirm earlier findings that students have difficulties understanding the nature of science.[18] Many high school students consider science as a collection of facts[19] and do not differentiate between observational evidence and explanations of this evidence.[20] Their views of the role of scientific models are contrary to the views of scientists,[21] and after years of formal schooling they lack an understanding of the main features that distinguish science from other mental enterprises.[22]

III. DESIGN OF THE STUDY

A. Procedure and course description

A student sample was chosen from the two-quarter physics sequence for participants in the Freshman Engineering Honors program at The Ohio State University. Any engineering first-year student identified by the university as an honors student could elect to take the course. The students in the program were generally very bright (their average ACT score is 29.83), and more than 97% of them had taken a year or more of physics in high school. The first 10-week quarter covered introductory mechanics; the second covered electricity and magnetism. Two instructors each taught a section of the course in the first quarter, and two different instructors each taught a section in the second quarter. Approximately two hundred students were enrolled in the sequence, evenly divided (or nearly so) between sections.

Research-based, active-learning strategies were employed extensively in all sections, and included interactive "lectures," cooperative group problem-solving, and other nontraditional innovations.[23] These methods had been used in this course for several years by one of the instructors. Each week students attend three 1-hour lectures, two 1-hour recitations, and one 2-hour laboratory, all employing these interactive engagement techniques most of the time. Recitations and laboratories (with about 30 students in each) were taught by teaching assistants (TAs). Course instructors and TAs met weekly to coordinate their instructional goals.

In addition, the course embodied a learning environment that mirrored the investigative character of science.[24] To construct physics concepts and laws, students first observed physical phenomena that were carefully selected by the instructor and presented in lecture,[25] then devised qualitative explanations for the patterns they observed and tested them by predicting the results of new experiments.[26] Then they developed mathematical explanations and designed experiments to test and find the limitations of these explanations, and finally applied concepts and skills to solve word and experimental problems (in recitation and laboratory). In this way, the course structure emphasized the process of science and the proper justification for scientific knowledge. From the epistemological point of view, the goal of the course was to help students construct knowledge following a possible path that a scientist might take, and thus help them replace their naïve epistemologies with the epistemology of physics.

Each week students reflected on what and how they learned by writing Weekly Reports. Students were asked to answer four open-ended questions:

(1) What did you learn in lab this week? How did you learn it?
(2) What did you learn in lecture and recitation this week? How did you learn it?
(3) What questions remained unclear?
(4) If you were the professor, what questions would you ask to determine if your students understood the material?

Students responded via the World Wide Web; responses were typically 1-page long, though many were longer.

Each week, half of the reports from each course section were randomly selected to be graded and given feedback. The first-quarter graders were a faculty member who had helped design the course (E.E.) and a graduate student; the second-quarter graders were the course instructors. Graders answered students' questions [from question (3)], and provided comments on questions (1), (2), and (4). The comments to questions (1) and (2) encouraged the students to be precise, clear, and complete in describing what and how they learned, and prompted them to refer to the in-class observations, experiments, and reasoning processes that were intended to help them learn physics. The graders' comments to question (4) emphasized the need for clarity and creativity. Weekly reports were worth 10% of the final course grade; as the students were told, points were deducted from each report for lack of clarity or thoroughness, but not for content.

B. Summative assessment instruments

We used three research-based multiple-choice instruments to assess aspects of the students' conceptual understanding and problem-solving ability. The Force Concept Inventory[27] (FCI) is a standard measure of students' understanding of Newton's laws of motion, and comprises 30 questions. It was given during the first quarter (on mechanics), pre- and post-instruction. The Mechanics Baseline Test[28] (MBT), a 26-item problem-solving measure for Newtonian mechanics, was given as part of the final exam for the first quarter. In the second quarter, the 32-item Conceptual Survey of Electricity and Magnetism[29] (CSEM) was given before instruction and again as part of the final exam.

Table I. FCI and CSEM normalized gains.

	Low gainers						High gainers					
Student	1	2	3	4	5	6	7	8	9	10	11	12
FCI	0.12	0.18	0.19	0.20	0.33	0.35	0.78	0.80	0.87	0.87	0.88	0.93
CSEM	0.06	0.05	0.31	0.09	0.25	0.22	0.72	0.60	0.64	0.67	0.80	0.77

C. The sample

From the half of the class that had the lower FCI pretest scores (that is, those who had the largest chance for gain), we chose the students with the lowest and the highest FCI normalized gains (about 10 of each). Normalized gain, also known as the Hake factor, is the raw gain divided by the maximum possible gain for a given pretest score.[30] We removed from the sample a few low-gaining students whose MBT scores were higher than the class average, recognizing the possibility that they had not taken seriously the FCI posttest (which was not part of their course grade). We also removed a few high-gaining students who did not achieve high gains on the CSEM, and one low-gaining student whose CSEM gain was very high. In this way, our sample ultimately consisted of six consistent ''high gainers'' and six ''low gainers,'' students who were either very successful in learning physics concepts in the course or very unsuccessful.

The normalized FCI gains and normalized CSEM gains for the students in each group are listed in Table I. As it turned out, the low gainers were from all course sections; the high gainers were all from one of the first-quarter sections and from both second-quarter sections. Thus, although the instructor may have had some effect in the first quarter, the effect disappeared in the second quarter.

D. Weekly reports

We observed students' reflective skills and epistemological views in the text of their reports. We collected 13–19 reports from each student in our sample. (Reports from the first two weeks of the first quarter were not available for analysis; thus, week 1 is actually the third week, week 2 the fourth, and so on to week 9. Numbering continues through the entire second quarter with weeks 10–19.)

Hammer's interviews of introductory physics students[16] focused on problem solving and textbook interpretation, and as such were domain specific. The same is true of the Weekly Reports; the open-ended questions posed each week, ''What did you learn?'' and ''How did you learn it?,'' ask students to address specifically their experience in the classroom.

Reports also represent a sufficiently different context from interviews to merit their use. Reports give an alternative perspective on student beliefs that may serve to replicate and verify the findings of interviews. Or, if they provide us with different information, then we can begin to map out the context dependence of epistemological beliefs. Also, while the context of Weekly Reports is not much more natural than that of interviews, they at least are an integral part of the course's instructional agenda. They have the advantage of being automatically generated from students' work in the course, and avoid problems with interpreting students' body language and tone of voice.

As one of us (E.E.) helped grade the first-quarter reports, she was familiar with the kinds of responses students wrote. We both read many reports of students not in the present sample and developed an initial coding scheme. The scheme was finalized after reading the reports of our sample students.

Fourteen codes were developed and set into three categories. We coded indications of *what* the students said they learned, *how* they said they learned it, and *inferences* we could make about other views about the nature of physics knowledge. Because the coding scheme itself is a product of our research, it is described in the next section.

To measure how articulate the students were about how they learned, total numbers of *code indications* were tallied. Code indications are instances of assigning a particular code to a sentence, a group of sentences, or an idea in a report. Information about each student's epistemological preferences was partly determined by normalizing the number of indications for each code. The normalization was calculated by dividing the number of indications for a particular code by the total number of code indications for that student. This method controls for how much students write about their learning. High and low gainers' normalized numbers of indications were compared for individual codes. Raw numbers were compared for particular code clusters, as described in Sec. IV.

E. Coding scheme for weekly reports

After reading the reports of several students, it was clear that they described learning in several different ways. By using our knowledge of epistemological dimensions identified by others[4,8,16] and our understanding of the epistemological goals of the course (constructing knowledge by replicating the processes of science), we were able to create a coding scheme for identifying and categorizing students' ideas. Examples of indications of each code from actual student reports are listed in the Appendix.

1. What they say they learned

Students usually listed the things they learned in a given week, making them relatively easy to identify. We considered each mention of something learned as evidence that the student thought it was important enough to mention in the report, if not actually important to learn. We used four codes in this category.

(1) *Formula*—equations or other mathematical statements, or the implication that they think formulas are important, without elaboration on their underlying meaning.
(2) *Vocabulary*—definitions or other physics language conventions.
(3) *Concept*—qualitative descriptions or mentions of concepts, ideas, relationships, or limitations of these.
(4) *Skill*—laboratory design skills, measurement skills, or problem-solving methods and skills, or the implication that they think skills are important.

2. How they say they learned

The ways in which students described how they learn are numerous. In reading the reports, we looked for indications of events that convinced the students that something was true. As expected, we found mentions of the direct transmission of information from authority (instructors or textbooks), and of more independent reasoning processes for gaining knowledge. Many students also described practice and simple observation as ways of learning. Some also mentioned or implied the role of prediction and testing in constructing understanding, an explicit focus of the course. We ultimately defined eight codes that describe indications of how students say they learn.

(5) *Observed phenomenon*—observed a physical phenomenon, demonstration, or experiment, without mention of what was learned in the process.
(6) *Constructed concept from observation*—learned a concept simply by observing a phenomenon, demonstration, or experiment (confusing an inference with an observation).
(7) *Reasoned/derived in lecture*—followed the reasoning process by which the large class came to a concept or formula, by using prior knowledge and experience, experimental data, logic, mathematics, and/or analogies.
(8) *Reasoned/derived in lab*—actively reasoned by one self or in a small group to come to a concept or formula, by using prior knowledge and experience, experimental data, logic, mathematics, and/or analogies.
(9) *Learned by doing*—learned a concept, definition, or formula by using it, or learned a skill or process by performing or practicing it.
(10) *Authority*—told or convinced by instructor, friend, textbook, or other authority figure.
(11) *Predicted/tested*—predicted the outcome of an experiment and then conducted or observed the experiment.
(12) *Predicted/tested/interpreted*—conducted or observed an experiment to test an idea and interpreted the results of that test.

3. Inferences about their views

Many statements made by students imply certain beliefs about the nature of physics knowledge. Several students mentioned the usefulness of physics knowledge in solving practical problems, and some expressed the expectation that physics knowledge should ''make sense'' or ''fit together'' coherently. These indications led to two more codes.

(13) *Applicability of knowledge*—indication of belief that physical laws or concepts can and should be applied to solve new problems.
(14) *Concern for coherence*—indication of belief that physical laws and concepts fit together into a coherent whole, or at least should agree with each other and with common sense.

4. Inter-rater reliability

After developing the above coding scheme, we performed a reliability check. The two of us independently coded the reports of four different students from the first quarter. In every single instance (sentence or group of sentences) in the reports, we agreed on which codes were indicated, although not always on the exact number of indications of each code. On that number we agreed 90% of the time.

5. Favorable and unfavorable codes

Because particular ways of learning were emphasized in the course, some code indications were deemed more appropriate than others. In this study, we identify epistemologically favorable codes as those that indicate student reflection on the construction of their own knowledge: reasoning using observational data or prior knowledge, experimental testing of ideas, concern for coherence [codes (7), (8), (12), (14)]. We call epistemologically unfavorable the codes that indicate that a student reported observations without mentioning making inferences, relied unduly on authority as a source of knowledge, or described testing experiments without reasoning or interpretation [codes (5), (10), (11)]. As described, ways of learning characterized by the favorable codes were emphasized in the course, and are thus considered appropriate subjects of reflection. Unfavorable codes represent ways of learning that are counter to the goals of the course. Although it is sometimes appropriate to learn some things by authority, this was rarely the case in this course.[31]

In addition to looking at codes individually, we summed the code indications about learning that were almost always appropriate in this course [favorable codes (7), (8), (12), and (14)] and those that were almost never appropriate [unfavorable codes (5), (10), and (11)]. The appropriateness of the other codes about how learning happens depended on the context of what was being learned.

IV. RESULTS

A. Descriptions of students' weekly reports

To provide a better understanding of the nature of students' reflection on learning, we describe here the reports of four students chosen because their reports demonstrate the diversity in the reports of the entire sample. Morris, a low gainer, did not write very much, and what he did write is mostly coded as unfavorable. Walter, another low gainer, also reflected on unfavorable ways of learning, but wrote a great deal. The reports of the two high gainers, Theo and Miles, show a large amount of reflection, but are focused on different, favorably coded ways of learning.

1. Morris (student 4)

Morris's reports were always very brief, receiving an average of only 3.3 codes per week. He was inarticulate, and usually mentioned only what happened in class rather than what and how he learned. Sometimes it was difficult to judge if his physics statements were correct because they were so poorly worded.

When Morris did address his learning, he usually did it by listing things that he learned by doing, mostly problem-solving skills. (In fact, of just 23 codings about how he learned, there were 14 indications of *learned by doing*.) He used the word ''tested'' occasionally, but did not say if there was any prediction; he may have used this word just because it was used in class. He described some experiments, but again in light of learning by doing rather than for observations or testing. A typical example from the first quarter is below.

(1) *What did you learn in lab this week*?

We learned that an object may have many components acting on it. There is always a vertical and horizontal component on an object in motion.

How did you learn it?

We constructed a track and raced ''Hot Wheels® car'' through the track. At the end of the track was a jump that propelled the car upward. This demonstrates that the horizontal component will tell how far the car will travel in a given time. The vertical component will determine the time the car spends in the air.

From the second quarter

(2) *What did you learn in lecture and recitation this week*?

We attempted to use of anti-derivitive (*sic*) formula to find the electrostatic potential.

How did you learn it?

With attempted derivitive (*sic*) problems drawn on the board.

2. Walter (student 5)

Walter was the most verbose low gainer, and wrote more than three times as much as Morris. He mentioned learning problem-solving and laboratory skills more often than anything else, and almost always indicated having learned by doing [code (9)] or from authority [code (10)]. In this typical example from the first quarter, he mentions the ActivPhysics simulation software[32] that was often used in lectures.

How did you learn it?

When we were learning about the projectile motion in class this week, the best way for me to understand the problems was to see [Instructor A] working on them via ActivePhysics (*sic*) on the board. The ActivPhysics is very helpful in portraying what actually happens in the experiment. It does not totally help though because it does not give you the equation. This is where [Instructor A] comes in. He uses the ''known'' quantities to find an equation to use and then uses it on the board to help find the answer that ActivPhysics gives us.

By [Instructor A] showing the experiments in class and working out the problems, I have been able to grasp some of the things we are learning so far.

It's clear that Walter was depending much more on the instructor than on himself, unlike most of the high-gain students.

However, there are indications that Walter's epistemological reflection was beginning to change. By the end of the second quarter, he referred less frequently to authority and paid more attention to his own role in learning. For example, he wrote:

How did you learn it?

We learned about electric fields by doing problems in class. We looked at electric fields in lecture and [Instructor B] taught us how they worked but what we did in recitation helped me the most. We did ... problems that helped us process what [Instructor B] was teaching us. By us doing the problems in groups nonetheless, it helped all of us figure out what needed to be done to determine how electric field affects point charges.

Also at this time, he began to indicate that he was looking for coherence in what he learns. None of the other low gainers showed such a change.

3. Theo (student 11)

This high-gain student did not write very much, but was very articulate about what he learned and showed a great deal of thought and personal involvement in learning. He often described his attempts to visualize equations and principles, and to integrate them with his existing knowledge. He used derivations and invented analogies, and always looked for cause-effect relationships in equations. This example, from the first quarter, shows his typical search for connections.

How did you learn it?

I learned about f=ma by understanding that if you touch something it starts moving. The kinematics equations alone do not explain this. It makes sense to have a component based on mass and acceleration. The bigger the object is the harder it is to push, and the harder you push the faster it goes.

Another example shows Theo's efforts to use his prior knowledge to help construct new knowledge:

(1) *What did you learn in lab this week*?

In lab we learned how to determine the x and y components of a Hot Wheels® car traveling in a form of projectile motion. We figured how to determine the velocity and angles without directly measuring them.

How did you learn it?

We learned this by applying the different kinematics and mathematical equations we learned. We had a starting point and taught ourselves how to get to the end point by just using what we already know.

Theo had a great willingness to think deeply about what he was learning and to extend it to new situations. This example shows that he was aware of the usefulness of analogies for developing understanding and knew that they have limitations.

(2) *What questions remained unclear*?

What causes gravity in general? Is gravity like a magnetic attraction? If so, then is it possible to counter it?

In the second quarter, Theo displayed even more reflection on aspects of knowledge and learning, especially the development of knowledge. When listing what he learned in a given week, he began to mention epistemological skills, such as ''how to prove'' a particular physics principle. As early as the first week of the second quarter, he responded with some interesting questions:

(3) *What questions remained unclear*?

Why was it decided to make the coulomb such a large unit?

Table II. Total number of code indications for each student.

	Low gainers						High gainers					
Student	1	2	3	4	5	6	7	8	9	10	11	12
No. of indications	94	94	169	43	216	119	236	239	181	202	134	152

(4) *If you were the professor, what questions would you ask to determine that your students understood the material*?

> What steps were taken to go from having no knowledge to having a general understanding of charged particles?

4. Miles (student 10)

Like Theo, Miles is a high gainer and understands physics very well. However, his reports are very different from Theo's. Miles was more articulate; he itemized what he learned and then meticulously addressed how he learned each item. Rather than focusing on his own role in constructing understanding, he addressed how the class developed knowledge and how knowledge was ultimately justified. Also unlike Theo, Miles never asked questions in part 3, and asked only unimaginative questions in part 4.

Miles recognized derivation as a method of acquiring knowledge, and was very adept at following the reasoning processes of the class, as this example shows.

(1) *What did you learn in lecture and recitation this week*?

> We learned about projectile motion. We learned that in projectile motion, the horizontal component of the velocity is always constant (1). And the total time in the air is determined solely by the vertical component of the velocity (2). So the vertical and horizontal components of the projectile velocity are independent. Also Newton's second law can be used to solve projectile motion problems (3). We also began to learn about circular motion. In circular motion the force is directed towards the center of the circle (4) and the object in circular motion will want to go in a straight line (5).

How did you learn it?

> We learned (1) through a simple demonstration where a cart with a vertical ball launcher, launched a ball by a sensor just before it went under a bridge and when the cart came out at the other side of the bridge it was caught in the holder. So the horizontal v was the same as the cart, which was constant, so it was too constant. We proved (2) by doing a problem where a canon (*sic*) is at a given angle and velocity. Then we use the vertical part of the v and found the time and used the time to find the horizontal displacement and that gave the right value. (3) was verified by being given the maximum vertical and horizontal displacements of a projectile. From there we found the angle to aim the canon (*sic*) to get the ball into the box. Then we tested our predicted value in the experiment and it worked. For the circular motion (4), we showed that by rolling a bowling ball and then hitting the ball with a hammer, the force, in different uniform directions in order to see which produced circular motion. And hitting towards the center was the one that worked. (5) was proved by having a circular track with a break at a point to see where the ball that was traveling in the circular motion would tend to go without the track and it showed that it would go straight from the last point in the circle.

Right from the beginning, Miles emphasized making predictions, testing them, and interpreting the results [as indicated by code (12)]. For example, he saw when test results did not make sense, and elaborated on possible causes. This emphasis continued throughout the two quarters. He also realized that the application of new knowledge to solving problems and the need for it to fit coherently with existing knowledge could serve as additional tests of its validity. This example refers to an experiment conducted by his lab group.

(2) *What did you learn in lab this week*?

> We learned the difference between static and kinetic friction and how to measure both types of friction.

How did you learn it?

> We found the coefficient of kinetic friction by adjusting the incline of the track so that the monster truck would neither move up nor backwards rather slide, that way it would be kinetic friction. Then with the appropriate force diagram we were able to calculate the coefficient. Then using our experimental value we predicted what hanging weight the truck could pull over a pulley on a lower incline. The prediction and the actual value were within 3 g so it was sufficient proof. To find the static friction we did the same method as for kinetic expect (*sic*) the truck was not in motion this time. To verify that value for static friction, we predicted what weight the nonmoving truck could start to move, and that value too verified our predictions.

Miles's reports did not change significantly in the second quarter, but remained excellent examples of very thorough reflection on the construction of new knowledge.

B. Amount of reflection of the two groups

The total number of codes attributed to the reports of each student (over the 20-week period) is listed in Table II. The high gainers clearly wrote (on average) more about what and how they learned than the low gainers. However, how much they wrote about their learning is not the whole story; *what* they wrote is just as important, as closer examination reveals.

Walter (student 5), the verbose low gainer with 216 code indications, frequently focused on authority as a source of knowledge, as described in more detail above. In fact, his weekly reports were coded for *authority* more than twice as much as the reports of any high-gain student. For example,

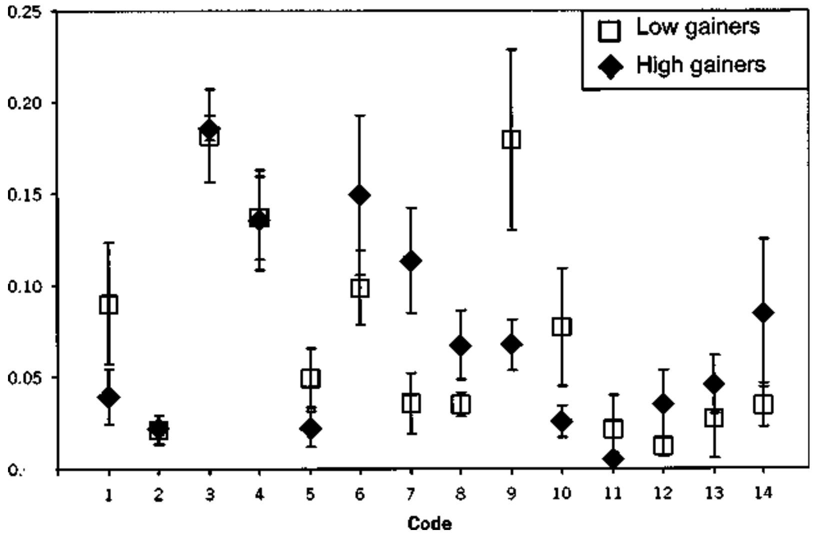

Fig. 1. Average normalized number of indications for each code, for high gainers and low gainers. The error bars are one standard deviation in length.

one week Walter wrote, ''We were told that an external force acting on a moving object can change the energy of the system.'' Meanwhile, the high gainers who wrote relatively little about how they learned (with 134 and 152 code indications) concentrated on their own personal role in constructing knowledge. As Theo (student 11) reported about a lab activity, ''We had a starting point and taught ourselves how to get to the end point by just using what we already know.''

C. Epistemological preferences

Although each student was somewhat consistent in what he or she wrote, clearly preferring to describe some ways of learning over others, each showed a number of different indications (that is, each student's collection of reports for the 20-week period merited several different codes). High and low gainers, on average, showed differences in some of the codes and not in others. The average normalized number of indications for each code for the two groups is shown in Fig. 1.

In terms of what the students say they learned [codes (1)–(4)], both groups frequently mentioned concepts and skills and hardly ever mentioned scientific vocabulary. However, the low-gain students mentioned equations and formulas [code (1)] much more often than the high gainers.

There were also important differences in how they said they learned. Codes (7) and (8) indicate that high gainers are much more focused on reasoning as a way of knowing than are low gainers, who mention experiments and observations without explanation [code (5)] more frequently than high gainers. Low gainers also indicate learning by doing [code (9)] much more often than the high gainers. Low gainers as a whole (not just the verbose student, Walter) have a stronger focus on authority [code (10)] than high-gain students. Making predictions, testing them, and interpreting the results was an important emphasis in the course. The related codes [(11) and (12)] did not appear particularly often, but still indicate an important difference between the groups. High gainers mentioned prediction and testing 52 times in all, and went on to describe interpretation of the results on 45 of those occasions. Those low gainers who mentioned prediction and testing, on the other hand, mentioned interpretation only 10 of 32 times. There was also a large difference in code (14), *concern for coherence*. The high-gain students much more frequently tried to connect what they were learning to their prior knowledge of physics or the natural world.

The raw numbers of favorable and unfavorable code indications for each student are shown in Fig. 2. It shows that the high gainers and low gainers focus on different aspects while reflecting on how they learned something. High gainers reflect more on the construction of coherent knowledge than low gainers, who focus more on rote learning. An indepen-

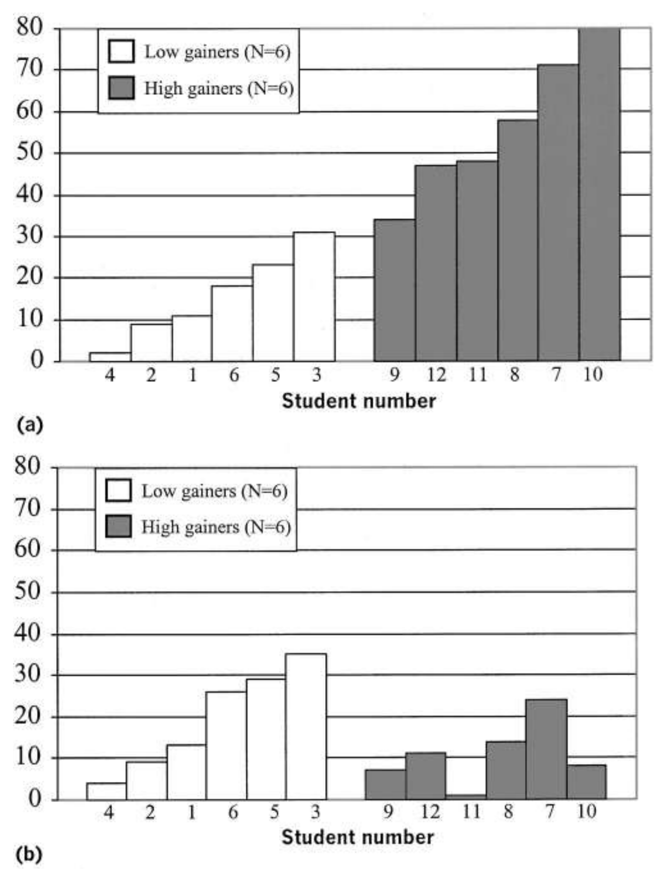

Fig. 2. Raw numbers of favorable (a) and unfavorable (b) code indications for low gainers and high gainers. In (a) numbers are ordered from smallest to largest within each group; in (b) they are arranged to match the student order of (a).

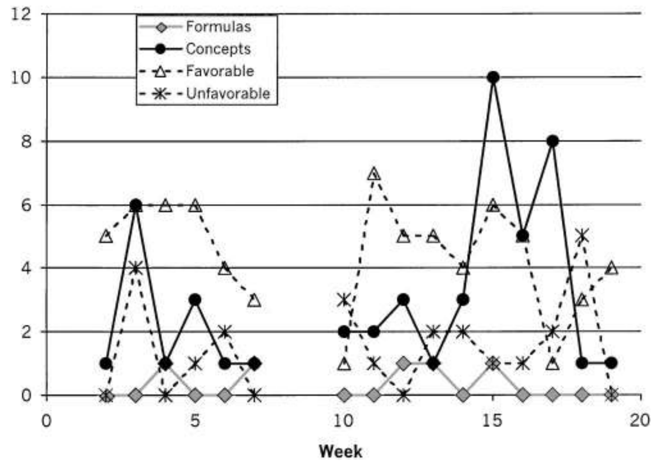

Fig. 3. Raw number of code indications per week for Gloria (student 7), including codes (1) (formulas) and (3) (concepts) and favorable and unfavorable code groups.

dent samples t-test shows the differences between high and low gainers' averages on each of these two composite measures to be significant at the $p < 0.05$ level.

The distribution of codes for different weeks shows that the character of student reflection on learning depends somewhat on the week, showing some dependence on the specific physics domain. Figure 3 shows this distribution for one high-gain student, Gloria (student 7). We can see that although Gloria usually focused on reasoning and other favorable learning approaches, there were more unfavorable reflection indications during weeks 3, 10, and 18. Instruction in those weeks focused on projectile and circular motion, electrostatic forces and Coulomb's law, and magnetic flux and electromagnetic induction, respectively. Gloria appears to have had difficulties interpreting numerous experiments that were shown in lectures during these weeks. An unsteady ''zigzag'' pattern is common to all students in the sample, for all codes.

V. SUMMARY AND CONCLUSIONS

The results of our analysis of Weekly Reports using our coding scheme suggest that different students do in fact reflect differently on the construction of knowledge in the same instructional environment. For example, even if the course is structured in an epistemologically favorable way and students do not receive new concepts from authority, some of them still think that they learn from authority.

We selected a sample of students for the study based on the improvement of their conceptual knowledge. All 12 students in the study started learning mechanics and electromagnetism in college with low pretest scores, compared with their classmates. Six of the students finished with low posttest scores in each quarter and six finished with high scores. This separation showed that students in the first group were less successful in learning in the course. We found that different students from two sample groups (low gainers and high gainers) exhibit preferences for different ways of learning. Our study addressed the specific research questions posed in Sec. I.

(1) What do students prefer to describe while reflecting on their learning, that is, what are their ''epistemological preferences?''

By analyzing students' answers to the question, ''How did you learn it?'' we found that students focused almost exclusively on experimental evidence, logical reasoning, practice, and authority. They also indicated common sense, the applicability of knowledge, and its coherence as factors affecting their learning. This focus allowed us to develop 14 codes for characterizing three aspects of student reflection (what they learned, how they learned it, and inferences about beliefs). This coding scheme is a representation of our findings.

The fact that it was possible to code student reflection on the construction of knowledge with a limited number of codes shows a general consistency across different students and diverse physics content. This finding suggests the possibility of using the same coding scheme to analyze student interviews and classroom interactions to compare findings of different studies.

(2) How might the amount of reflection they exhibit relate to their conceptual learning gains?

Our results suggest a correlation between students' conceptual gains and their ability to reflect on their learning. Most of our low gainers did not write much about how they learned, compared with the high gainers. At the same time we found exceptions to this rule: in our small sample, one low gainer reflected on his learning in great detail, while two high gainers did not write a great deal. These high gainers nonetheless were able to reflect on the construction of knowledge by following the reasoning process in class or by making knowledge relevant to their personal experience. They also tried to make coherent sense of the material by asking profound questions. These exceptions might mean that it may not be the quantity of reflection, but its quality that matters, and student questions might provide worthwhile insights into their epistemological preferences.

(3) How might their epistemological preferences relate to their conceptual learning gains?

Analysis of specific codes suggests a possible correlation between conceptual gains and epistemological views. Low conceptual gainers were more likely than others to mention learning activities that are epistemologically less desirable: learning formulas without heeding their conceptual implications, learning from authority, and predicting and testing without interpretation. High gainers, however, more frequently referred to reasoning and interpretation of experimental results, and showed more concern for the coherence of knowledge than their counterparts.

More extensive research is needed to verify these tentative relationships. However, the possibility of such connections implies that ''good'' students have knowledge that is appropriate epistemologically as well as conceptually, and that they are better at reflecting on what they learn and how they learn it.

VI. IMPLICATIONS FOR INSTRUCTION

Our results imply that when it comes to learning physics concepts, student epistemologies matter. This conclusion is consistent with the research of others in different contexts.[6,8,16] It is only the beginning of our investigation of this connection, but this study of 12 students during 20 weeks of college physics instruction suggests that we might be able to enhance student content learning by encouraging appropriate epistemologies. If this is in fact the case, then there are strong implications for instruction. Although the effects of the approaches listed below have not been carefully measured, we have used them in this course and recognize that they might contribute to the development of sophisticated epistemological thinking:

Students could be encouraged to reflect on a regular basis on how they construct content knowledge and acquire skills. Although Weekly Reports are a time consuming way to do it, the same goal can be achieved by putting similar questions in homework assignments or in laboratory reports. The latter approach also encourages students to see reflection as an integral part of doing science.

Another way to encourage content-based reflection might be to ask content questions that call for the justification of knowledge, such as, ''How can you convince a friend that two objects always act on each other with forces that are equal in magnitude?''

We can also design questions that will indirectly encourage students to reflect on how they know what they know by asking them to make a decision, for example, ''You have a motorized toy car. How can you find out if it moves with constant velocity, constant acceleration, or changing acceleration?''

These are examples of open-ended questions that do not have a single solution. In addition to developing student epistemologies that might enhance student conceptual learning, they promote metacognitive (self-reflection and monitoring) skills.[11] These are high-level thinking skills that do not develop if students only solve problems with a known answer. More research is needed to link together these three aspects of learning: content acquisition, epistemology, and self-reflection. Do we want to develop epistemology and self-reflection to enhance content learning, or should learning of content be viewed as a vehicle to develop student epistemologies and high-order thinking skills? Current demands of the workplace for investigative and problem-solving skills[33] suggest that for college graduates, the second choice is also very important.

ACKNOWLEDGMENTS

This research was supported in part by NSF Grant No. GER-9553460. We are grateful for the support and suggestions of David Mills from Monash University, and for the numerous helpful comments of Alan Van Heuvelen, Gordon Aubrecht, and other members of the Physics Education Research Group at The Ohio State University. We would also like to thank Jonathan Pelz, Kathleen A. Harper, and Andrew Heckler for their support and cooperation.

APPENDIX: EXAMPLES OF CODE INDICATIONS

Each example is taken directly from students' reports, with no modification.

What they say they learned

(1) *Formula*—We learned Newton's third law where F (1 on 2) = −F (2 on 1).

(2) *Vocabulary*—We learned that units of power are called watts.

(3) *Concept*—We learned that when the sum of forces acting on the object is not zero, there is an acceleration of the object but when they are in the equilibrium the object moves at constant velocity.

(4) *Skill*—We learned that when dealing with a complex force problem, splitting forces into components is a way to solve it.

How they say they learned

(5) *Observed phenomenon*—We observed that the insulation pipes rubbed with natural fur repel each other, if one pipe is rubbed with the natural fur and the other one with synthetic, they attract each other.

(6) *Constructed concept from observation*—Then we observed a ball being compressed on a spring and watching the spring shoot the ball up. This displayed the elastic potential energy of a spring.

(7) *Reasoned/derived in lecture*—We derived the expression $v = ir$ with simple experiments where a certain current was placed into a circuit and compared with voltage. We then found a liner relationship between them and found the equation.

(8) *Reasoned/derived in lab*—We then determined the wiring of a box with six light bulbs and six switches which had up to three positions. We did this by applying what we know about the properties of loads wired in series and in parallel and observing the circuit's behavior under different combinations of switches.

(9) *Learned by doing*—We did context rich problems in recitation to practice dealing with the properties of circuits.

(10) *Authority*—The professor gave us the equation for the law of gravitation.

(11) *Predicted/tested*—We used the range equation to predict where the ball will land but it landed short.

(12) *Predicted/tested/interpreted*—We learned that Newton's second law can be used in combination with kinematics equations. For this we constructed an experiment and made a prediction based on the laws and equations and then found an experimental value. There were two hanging weights (700 g and 500 g) connected by a string across two pulleys. We derived how long it takes 700 g mass to hit the ground. We calculated it to be 1.31 s and got the experimental value to be 1.32 s which is close enough to verify that we can combine Newton's laws with kinematics.

Inferences about their views

(13) *Applicability of knowledge*—We built a horizontal and a vertical accelerometer. The accelerometers were another application of Newton's second law.

(14) *Concern for coherence*—In order to understand, why there are two different equations for gravitational potential energy we derived the simple mgy for close to the surface from the other equation.

[1]For example, *Webster's New Collegiate Dictionary* (G. & C. Meriam Company, Springfield, MA, 1974), p. 385, defines epistemology as ''the study or a theory of the nature and grounds of knowledge, especially with reference to its limits and validity.''

[2]For example, there were six talks related to epistemological issues given at the 2001 Winter Meeting of the American Association of Physics Teachers, 8 at the 2001 Summer Meeting, and 17 at the 2002 Winter Meeting; note that these figures do not include talks given at the Physics Education Research Conference that followed the summer meeting.

[3]E. Etkina, ''Weekly Reports: A two-way feedback tool,'' Sci. Educ. **84**, 594–605 (2000). David Mills suggested additional modifications to the reports.

[4]B. Hofer, ''Personal epistemology research: implications for learning and

instruction,'' Annual Meeting (American Educational Research Association, Seattle, WA, 2001).

[5]W. G. Perry, *Forms of Intellectual and Ethical Development in the College Years: A Scheme* (Holt, Rinehart and Winston, New York, 1970); M. F. 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); M. B. Baxter Magolda, *Knowing and Reasoning in College: Gender-Related Patterns in Students' Intellectual Development* (Jossey Bass, San Francisco, 1992); 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, 1994); D. Kuhn, *The Skills of Argument* (Cambridge University Press, Cambridge, UK, 1991).

[6]Marlene Schommer, ''The effects of beliefs about the nature of knowledge on comprehension,'' J. Educ. Psychol. **82**, 498–504 (1990).

[7]M. Schommer, A. Crouse, and N. Rhodes, ''Epistemological beliefs and mathematical text comprehension: Believing it is simple does not make it so,'' J. Educ. Psychol. **84**, 435–443 (1992).

[8]Barbara K. Hofer and Paul R. Pintrich, ''The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning,'' Rev. Ed. Res. **67**, 88–140 (1997).

[9]G. Qian and D. Alvermann, ''Relationship between epistemological beliefs and conceptual change learning,'' Reading & Writing Quarterly **16**, 59–74 (2000).

[10]D. Kuhn, ''Theory of mind, metacognition, and reasoning: A life-span perspective,'' in *Children's Reasoning and the Mind*, edited by P. Mitchell and K. J. Riggs (Psychology Press, Hove, UK, 2000), pp. 301–326.

[11]K. S. Kitchener, ''Cognition, metacognition, and epistemic cognition,'' Hum. Dev. **26**, 222–232 (1983).

[12]David M. Hammer and Andrew 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 (Erlbaum, Mahwah, NJ, 2002).

[13]S. S. Stodolsky, S. Salk, and B. Glaessner, ''Student views about learning math and social studies,'' Am. Ed. Res. J. **28**, 89–116 (1991); Barbara K. Hofer and Paul R. Pintrich, ''Disciplinary ways of knowing: epistemological beliefs in science and psychology,'' Annual Meeting (American Educational Research Association, Chicago, IL, 1997).

[14]J. Leach, R. Millar, J. Ryder, and M.-G. Séré, ''An investigation of high school and university science majors' epistemological reasoning in the context of empirical investigations,'' Annual Meeting (American Educational Research Association, Montréal, PQ, 1999).

[15]Andrew Elby, ''Helping physics students learn how to learn,'' Am. J. Phys. **69**, Suppl. 1 S54–S64 (2001); David M. Hammer, ''Two approaches to learning physics,'' Phys. Teach. **27**, 664–670 (1989); W.-M. Roth and A. Roychoudhury, ''Physics students' epistemologies and views about knowing and learning,'' J. Res. Sci. Teach. **31**, 5–30 (1994).

[16]David M. Hammer, ''Epistemological beliefs in introductory physics,'' Cogn. Instruct. **12**, 151–183 (1994).

[17]Edward F. Redish, Jeffrey M. Saul, and Richard N. Steinberg, ''Student expectations in introductory physics,'' Am. J. Phys. **66**, 212–224 (1998).

[18]N. B. Songer and M. C. Linn, ''How do students' views of science influence knowledge integration?'' J. Res. Sci. Teach. **28**, 761–785 (1991); F. Abd-El-Khalick, R. L. Bell, and N. G. Lederman, '' The nature of science and instructional practice: making the unnatural natural,'' Sci. Educ. **82**, 417–436 (1998).

[19]American Association for the Advancement of Science, *Science for All Americans: A Project 2061 Report on Literacy Goals in Science* (American Association for the Advancement of Science, Inc., Washington, DC, 1989).

[20]J. Biggs and K. Collis, *Evaluating the Quality of Learning: The SOLO Taxonomy* (Academic, New York, 1982); J. Clement, ''Overcoming students' misconceptions in physics: The role of anchoring intuitions and analogical validity,'' in *Proceedings of the Second International Seminar on Misconceptions and Educational Strategies in Science and Mathematics*, edited by J. Novak (Cornell University, Ithaca, NY, 1987), Vol. 3, pp. 84–97.

[21]L. Grosslight, C. Unger, E. Jay, and C. L. Smith, ''Understanding models and their use in science: Conception of middle and high school students and experts,'' J. Res. Sci. Teach. **28**, 799–823 (1991).

[22]F. Reif and J. Larkin, ''Cognition in scientific and everyday domains: Comparisons and learning implications,'' J. Res. Sci. Teach. **28**, 733–761 (1991); K. Sullenger *et al.*, ''Culture wars in the classroom: Prospective teachers question science,'' *ibid.* **37**, 895–916 (2000).

[23]These are described in more detail by Alan Van Heuvelen and Kathleen M. Andre, ''Calculus-based physics and the engineering ABET 2000 criteria,'' APS Forum on Education newsletter, Spring/Summer 2000, pp. 5–6.

[24]E. Etkina, ''Can we use the processes of physics to guide physics instruction?'' published online at http://www.gse.rutgers.edu/people/ee.htm, 2001.

[25]Here we mean the selection of one or more observational experiments (really demonstrations) in which most students are able to see a clear pattern. For example, in a lecture where students construct the idea that for an object to move in circle with constant speed, there must be a net force pointed towards the center, students first observe the instructor rolling a bowling ball along a long table (the ball moves with constant velocity) and then tapping it in the direction of motion (the ball speeds up). Then the instructor repeats the experiment tapping the ball in the direction perpendicular to the original direction of motion (the ball follows a parabolic path). The third time, the instructor taps the ball in the direction perpendicular to the direction of motion at every point. Students are asked to construct free body diagrams for each situation and explain the motion of the ball.

[26]In these ways, this cycle is different from science learning cycles suggested by R. Karplus and C. Lavatelli, *The Developmental Theory of Piaget: Conservation* (Davidson Film Producers, San Francisco, 1969) or by A. E. Lawson, M. R. Abraham, and J. W. Renner, *A Theory of Instruction: Using the Learning Cycle to Teach Science Concepts and Thinking Skills* (NARST, Cincinnati, OH, 1989). The Lawson *et al.* cycle begins with a question, rather than with an observation that generates questions. In Karplus and Lavatelli's cycle, explanation construction is followed by application rather than testing.

[27]D. Hestenes, M. Wells, and G. Swackhamer, ''Force Concept Inventory,'' Phys. Teach. **30**, 141–158 (1992).

[28]D. Hestenes and M. Wells, ''A Mechanics Baseline Test,'' Phys. Teach. **30**, 159–166 (1992).

[29]D. Maloney, T. O'Kuma, C. Hieggelke, and A. Van Heuvelen, ''Surveying students' conceptual knowledge of electricity and magnetism,'' Am. J. Phys. **69**, Suppl. 1 S12–S23 (2001).

[30]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**, 64–74 (1998).

[31]In this class, the instructor made a special effort to avoid ''content delivery'' and to engage students in the construction of concepts (from observations or relationships to other concepts) and then by testing the concepts experimentally. Physical notions were not defined before students constructed their meaning. Mathematical relationships were either discovered as patterns in data or derived from previous relationships. All relationships were tested experimentally.

[32]Alan Van Heuvelen, *ActivPhysics 1* (Addison Wesley Interactive, New York, 1997).

[33]See, for example, the ABET 2000 criteria, http://www.abet.org/eac/2000.htm; *What Work Requires of Schools: A SCANS Report for America 2000* (U.S. Dept. of Labor, Washington, DC, 1991); D. Rosdil, *What are Masters Doing?* (AIP Statistics Div., College Park, MD, 1996), Publication No. R-398.1.