Correlations between Student Discussion Behavior, Attitudes, and Learning

Gerd Kortemeyer*
Lyman Briggs School of Science
Michigan State University
(Dated: January 4, 2007)

An important result of Physics Education Research is that students' learning and success in a course is correlated with their beliefs, attitudes, and expectations regarding physics. However, it is hard to assess these beliefs for individual students, and traditional survey instruments such as the Maryland Physics Expectations Survey (MPEX) are intended to evaluate the impact of one or more semesters of instruction on an overall class and improve teaching.

In this study, we investigate the possibility of using the analysis of online student discussion behavior as an indicator of an individual student's approach to physics. These discussions are not tainted by the effects of self-reporting, and are gathered in authentic non-research settings, where students attempt to solve problems in the way that they belief is most efficient and appropriate.

We calculate the correlation of both MPEX and student discussions with different measures of student learning, and find that on an individual base, student discussions are a stronger predictor of success than MPEX outcomes.

PACS numbers: 01.40.Fk

I. INTRODUCTION

In the study of student learning, epistemological beliefs are defined as beliefs and views about how knowledge is constructed and evaluated. Typical dimensions along which these are evaluated include measures of independence in learning (taking responsibility versus relying on authority (instructors, books)), coherence (actively attempting to integrate new knowledge into a coherent framework versus seeing each piece of knowledge as a standalone entity), and emphasis on concepts (e.g., attempting to understand formulas versus memorizing them).

Most physicists would readily acknowledge that these beliefs and views are important factors in the learning and application of physics, and oftentimes refer to "the physics-way of thinking about the world and solving problems." Unfortunately, these beliefs are not very tangible: it appears almost impossible to teach them, and extremely hard to measure and assess them.

Traditional ways of assessing these beliefs include observations, guided interviews, and the deployment of surveys. Observation techniques, for example of groups of students solving problems, are not very scalable due to the need of setting up recording equipment and subsequent transcription. Also, students are aware of being observed, and may act differently than in unobserved scenarios. Guided interviews share the same limitations in terms of scalability and are possibly even more strongly influenced by efforts to make a good impression or please the interviewer.

Surveys, on the other hand, are very scalable and thus used most frequently. Examples are the Maryland Physics Expectations Survey (MPEX) [1], the Epistemological Beliefs Assessment for Physical Science (EBAPS) [2], or the Colorado Learning Attitudes about Science Survey (CLASS) [3], or through structured interviews (see for example [4, 5]). While these instruments take different approaches and have different philosophies behind their designs, they do have in common that the students need to react to statements outside of the normal course activity, and that they need to self-report their responses.

The MPEX makes the limitations of this approach very explicit in their "Product Warning Label" [6]: "students often think that they function in one fashion and actually behave differently. For the diagnosis of the difficulties of individual students more detailed observation is required." Online student discussions associated with online physics problems are different in that they are generated within the real context of the course, and students have a vested interest in making these discussions as productive as possible, given their understanding of how physics is done and their approach to it. Online discussions however are still scalable, since the students self-transcribe their contributions, and they are happening on a large enough scale that the awareness of being observed wears off. They could thus be a "reality check" of students' beliefs, attitudes, and expectations.

The MPEX "Product Warning Label" continues, that "this survey is primarily intended to evaluate the impact of one or more semesters of instruction on an overall class" [6], and recommends using the outcomes, in combination with evaluations of student learning of content, as a means to improve overall course instruction. In this paper, we are asking the question if the evaluation of student online discussion behavior can be used as a means to assess the attitudes and beliefs of an individual student, and if in turn, these can be used to predict the success of an individual student in the learning of physics content. An obvious application would be the early detection of

---

*Electronic address: korte@lite.msu.edu; URL: http://www.lite.msu.edu/kortemeyer/

students at risk.

In particular:

- We classify the online homework discussion contributions from one course

- We deploy the MPEX for comparison as a pre- and post-test

- We are using the pre- and post-FCI, as well as the final exam and course grades, as a measure of student learning

- We correlate online discussion behavior and MPEX cluster scores with measures of student learning

II. BACKGROUND

For an expert physicist, physics is much more than a vast unconnected collection of facts or the mechanics of manipulating given formulas; instead, it is a way of thinking. Students, on the other hand, tend to work from a not yet coherent set of physics factoids organized by surface features rather than physics concepts [7] and often focus on finding the right formula for the situation rather than on the construction of new knowledge and relationships [8, 9]. These and other beliefs about what constitutes knowledge in physics and how one develops knowledge are described as epistemological beliefs [5, 10, 11].

Different epistemological beliefs can lead to very different understandings of the same scenario. For example, expert physicists see the derivation of a formula as a way to embed the new knowledge into their existing framework, while students tend to see it as a proof that a formula is true and "okay to be used" [1]. As a result, the whole process is frequently ineffective in lectures, since students would see a proof of correctness as irrelevant and would rather just trust the "authority"[31].

Previous studies indicate that correlations between epistemological beliefs and academic performance exist, both directly and indirectly [11, 12]. For example, Schommer [13] found that belief in "quick learning" (characterized by seeking single answers, avoiding ambiguity, and relying on authority) negatively correlates with the GPA of secondary students, even after controlling for general intelligence. May and Etkina [14] found possible correlations between epistemological beliefs extracted from extensive lab reports and conceptual learning gain in introductory physics courses. For example, students who stated that they learned formulas (rather than investigated their conceptual implications), relied on authority, and made no efforts to interpret results were found to have lower gains on the Force Concept Inventory, Mechanics Baseline Test, and Conceptual Survey of Electricity and Magnetism. It should be remarked that the performance on conceptual tests is not necessarily directly connected to good course grades, which is an issue

that students are aware of [15] and can lead them to act contrary to their beliefs.

Correlations between epistemologies and learning alone do not imply causal relationships. Based on videotaped classroom scenarios, written work, and interviews, Lising and Elby [16] argue that a more expert-like epistemology indeed leads to better learning, and thus, curricular materials and teaching techniques should explicitly attend to students' epistemological beliefs.

The problem is how to measure these beliefs and worsened by the fact that these beliefs might shift and that categorical approaches (i.e., dividing students into classes that hold or do not hold a certain beliefs) might be inappropriate [17], as well as the fact that the students do not necessarily act according to what they know to lead to deeper learning [15] . Techniques include surveys, guided interviews, and observations. While interviews and observations are likely resulting in better data, the effort in conducting them also limits the scale at which they can be conducted. Surveys do not have this scalability problem, but research results regarding the predictive power of these instruments are not always conclusive: for example, Coletta and Philips [18] found a strong correlation between the MPEX and FCI Gain, while Dancy [19] found low correlations between the MPEX and the performance on homework, tests, and final exams. It is unclear why these studies would come to such different results regarding the predictive power of the MPEX on an individual student level. Until more insights are gained, it remains a good idea to abide by the "Product Warning Lab" [6] that the survey is best used to gain insights into the beliefs of the class as a whole.

Online discussions take place within the regular course context and over its complete duration. They are a rich source of feedback to the instructor [20], and their quality and character was found to be correlated with the type and difficulty of the associated problems [21], i.e., data exists regarding the influence of *problem* characteristics on associated discussions. Unfortunately, less data exists on the correlation between *student* characteristics and discussion behavior, because usually only very few student characteristics are known, with the exception of the students' overall performance in the course. Thus, one of the few findings was the fact that certain discussion behavior, most prominently exhibited on "non-sanctioned" discussion sites external to the course, is negatively correlated with performance in the course [21, 22].

Few studies exist on the correlation between beliefs data gathered in research settings and actual discussion behavior in the course. For example, Hogan [4] assessed eight graders' epistemological frameworks through interviews and then analyzed their discussion behavior in a science course with a particular focus on collaboration, finding a number of correlations. In the interviews, students were ask to articulate views about themselves, about how they learn, and about the subject area. It was found that students' views on learning most strongly correlated with their peer-discussion behavior, for example,

students who exhibited a constructivist view of learning were also most strongly engaged in the peer-discussions and collaborative knowledge building. Hogan's study supports the importance of the distinction between public and personal epistemologies: public epistemologies are beliefs held about the general nature of discovery and knowledge in the scientific community, while personal epistemologies are beliefs about one's own knowledge and learning [12, 16]. Discussion behavior seems to be most closely related to personal epistemologies [4].

In this paper, we investigate if correlations similar to those found by Hogan exist between online peer-discussions and epistemological beliefs, and if those in turn are correlated with measures of student learning.

III. SETTING

The project was carried out in an introductory calculus-based physics course with initially 214 students. Most of the students in this course plan on pursuing a career in a medical field. The students are members of a residential college on campus of Michigan State University which stresses courses about the nature and philosophy of science (public epistemology). The course had three traditional lectures per week. It did not use a textbook, instead, all course materials were available online. Topics were introductory mechanics, as well as sound and thermodynamics. There was twice-weekly online homework: one small set as reading problems due before the topic was dealt with in class (implementing JiTT [23]), and a larger set of traditional end-of-the-chapter style homework at the end of each topic. The online problems in the course were randomized using the LON-CAPA system, i.e., different students would receive different versions of the same problem (different graphs, numbers, images, options, formulas, etc) [24, 25]. The students had weekly recitation sessions, and a traditional lab was offered in parallel. The course grade was determined from the students' performance on biweekly quizzes, the final exam, the recitation grades, and the homework performance.
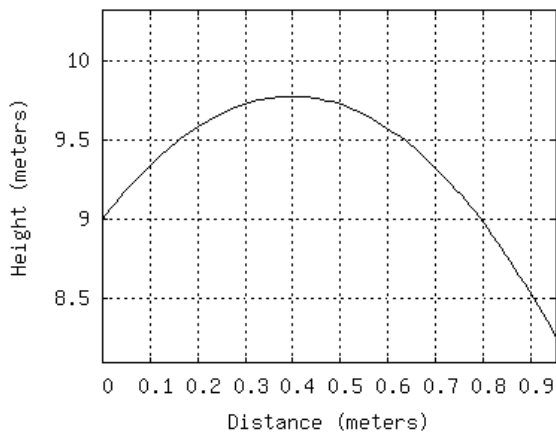
IV. MEASURES AND INSTRUMENTS

A. Discussion Analysis

In LON-CAPA, discussions boards are directly underneath each online homework problem, i.e., each problem has its own discussion board on the same web page. Since problems in LON-CAPA are randomized, such that each student has different options, numbers, graphs, equations, or even scenarios (e.g., accelerating to the left or to the right), the students cannot simply exchange the correct answer, and are encouraged to freely discuss the problems with each other. Figure 1 shows a screenshot of such a problem and its associated discussion.

The author analyzed the online student discussions that were associated with the online homework given in his course, using the scheme first suggested in Ref. [21]. The student names were not available during classification in order to avoid bias. There were a total of 2405 such online discussion contributions over the course of the semester, where one posting counts as one contribution. A particular contribution can receive more than one classification.

The following list shows the classifications taken into consideration, as well as illustrative examples that would receive the respective classification.

- Discussion contributions were classified as *surface* if they dealt with surface features of the problem or were surface level requests for help.

  - You use for the first part, c(ice)=.5, c(water)=1, Lfluid=80 right?

  - what do d and g stand for?

  - "e" for this equation is equal to one because it is a black body ... hope this helps.

- *Procedural* contributions describe or inquire about mechanisms of solving a problem without mentioning the underlying concepts or reasoning.

  - ok in this problem make sure your temp. is in K and your c is in J. So for water use 4186 and for Alcohol use 2430. The problem is set up the same way that Mr. Anonymous has posted above. Hope that helps

  - ok, my original angle is 67 degrees. When i seperate the strings into their two right-angle triangles, the angle becomes 33.5 with the adjacent side being 43.7N (my force of gravity). to solve for the string, or my hypotenuse, i used h=43.7N/cos(33.5)= 52.41... then i tried to plug all of this into T= 52.41/2cos (67)=10.23 and that was wrong so then i did T=52.41/2cos(33.5) = 31.4 which was also wrong... why am i so wrong?! please help!!!

- *Conceptual* contributions deal with the underlying concepts of the problem.

  - Just common sense.
  I didnt use math, I just sorta looked at how loud the intensities they gave were and compared it to other noises that were at the same intensity. For instance:
  a nuclear explosion is around 220 dB from 500 m away.

The plot shows the trajectory (height versus distance) of an object launched at an angle of 75.6 degrees. What was the initial speed of the object? **4.0 m/s**
Computer's answer now shown above. Tries 0/12

Threaded View Chronological View Sorting/Filtering options Export?

Anonymous 1 (Fri Sep 22 01:26:29 2006 (EDT))

any hints to start?

Re: *Anonymous 2* (Fri Sep 22 01:56:48 2006 (EDT))

You need to find the Y component of velocity... you can do this by finding the height traveled (notice it does not start on the ground) and combining that with acceleration in a kinematics equation. From there use trig to get the original velocity.

Re: Re: *Anonymous 1* (Fri Sep 22 12:10:37 2006 (EDT))

how can we find the height traveled and how can we get the acceleration if we don't have the time?

Anonymous 3 (Fri Sep 22 16:41:27 2006 (EDT))

i'm lost on this one... can anyone help?

Re: *Anonymous 4* (Fri Sep 22 20:02:45 2006 (EDT))

Use the squared kinematics equation - so Vf^2 = Vi^2 + 2a (Xf-Xi).

FIG. 1: Example of an online homework problem with associated student discussions. The problem is past its due date, so the correct answer is shown. If it were still open, the students would have one text answer box, into which they would enter the value with physical units. The students in this example chose to post anonymously, however, the student name is still available to faculty.

a rock concert is around 120 dB
a person shouting is 80 dB.
So look at the answers and think about the question logically.

- I am totally lost on this problem. Could someone try to explain how you would approach this problem. I just don't understand how Newton's third law applies and why we would completely ignore the mass of the truck and just use the mass of the car.

- I have the correct answer, but I don't understand why it is correct. Why would there be an acceleration at the ball's highest point? Why wouldn't it be zero?

• *Solution-oriented* – the goal of the contribution is to arrive at the correct answer without mentioning or dealing with the mathematics or physics of the problem.

- First you want to take your total mass of the car and the two people (In my case 800kg of car + 100kg of people) and divide it by 4.
Then plug that mass into the equation: w=(k/m)^.5 K is given to you in the problem.
Then plug w into the equation Period=2pi/w
And theres your answer. I hope that helps!

-      Use   this   formula:

T(final)=(m1c1T1+m2c2T2)/(m1c1+m2c2). Convert temp to Kelvin and then for your final answer convert back to Cel.

- *Mathematical* – the contribution deals mostly with the mathematical aspects of the problem.

  - Change your commas to plus signs. Its asking you for the scalar product which means you must add the product of each component.

  - What's an arctan?

- *Physics* – the contribution deals with the physics-related aspects of the problem.

  - the SI unit of Power is a Watt (W)

  - velocity=flow/area

  - Why does the value of G not depend on which planetary system you study? That doesn't make any sense to me. I thought the gravitational forces of different planets are different. That's why we float on the moon...

One particular contribution could receive more than one classification. Each contribution was weighed by its length when calculating the overall discussion behavior of an individual student. For additional details, see Ref. [21], where the above classes were referred to as "superclasses."

Note that correctness of the contribution was not considered. For example, in the last physics-related example, the fact that the student confused the gravitational constant and the gravitational acceleration was not taken into account.

Interrater reliability was assessed by asking a graduate student in physics education research at another university to read Ref. [21] and classify a randomly selected subset of 104 contributions. Without further training, the overall reliability was 81 percent. As it turned out, though, this value did vary by class: the conceptual and mathematical classes had an interrater reliability of 91 and 96 percent, respectively, while the solution-oriented class had a reliability of only 58 percent. Looking at the prominence of contributions in certain classes, the author classified 60 percent and the graduate student only 38 percent as "solution-oriented." The discrepancy might be due to the fact that the author in his course attempts to strongly discourage this behavior and thus may be more prone to detect and label it than an individual who is not connected with the course. The discrepancy could likely be resolved with training.

As already found in Ref. [21], most students are quite prolific in their online discussions, but a few students only made a small number of contributions, leading to small statistics on their actual discussion behavior. For each of the discussion correlations, we thus also carried out a second calculation limited to students who contributed at least five entries over the course of the semester.

B.   The MPEX

We deployed the Maryland Physics Expectations Survey (MPEX)[1] both at the beginning and the end of the mechanics semester. Participation was voluntary. We calculated the "score" in comparison to the "favorable" expert responses given in Ref. [1] – please note that the word "score" in the context of the MPEX is thus not an absolute measure of correctness, but of agreement with the majority of an expert group, who does not even necessarily agree among each other. We calculated the final (post) deployment score, as well as, for students who participated both times, the gain. The same analysis was done for each cluster of the MPEX (example statements are given, including the expert answer):

- *Independence:* student takes responsibility for constructing their own understanding, rather than takes what is given by authorities (teacher, materials) without evaluation

  Unfavorable: In this course, I do not expect to understand equations in an intuitive sense; they must just be taken as givens.

- *Coherence:* student believes that physics needs to be considered as a connected, consistent framework, rather than a set of unrelated facts or "pieces"

  Unfavorable: Knowledge in physics consists of many pieces of information each of which applies primarily to a specific situation.

- *Concepts:* student stresses the understanding of the underlying ideas and concepts, rather than the memorization and usage of formulas

  Favorable: When I solve most exam or homework problems, I explicitly think about the concepts that underlie the problem.

- *Reality Link:* student believes that the ideas learned in physics are relevant and useful in a wide variety of real contexts, rather than having little to do with experiences outside the classroom

  Unfavorable: Physical laws have little relation to what I experience in the real world.

- *Math Link:* student considers mathematics as a convenient way of representing physical phenomena, rather than viewing physics and mathematics as having little or no relationship

  > Unfavorable: All I learn from a derivation or proof of a formula is that the formula obtained is valid and that it is OK to use it in problems.

- *Effort:* student makes the effort to use the available information and tries to make sense of it

  > Favorable: I go over my class notes carefully to prepare for tests in this course.

The overall scores (i.e., agreement with the expert group) of the students on the MPEX clusters were low (Independence 42%; Coherence 46%; Concepts 48%; Reality Link 55%; Math Link 40%; Effort 47%).

C. Measures of Student Learning

As a measure of student conceptual understanding and learning, we deployed the revised Force Concept Inventory (FCI)[26] at the beginning and the end of the course, again with voluntary participation. As an additional measure of student performance, the performance on the final exam and the course grade for each student were taken into consideration. For the grade we used the raw percentage score, not the number grades, since it provides finer grained information about the overall student performance in the course.

V. STUDENT PERCEPTION OF THE ONLINE DISCUSSIONS AND SURVEY INSTRUMENTS

An additional survey was deployed online after the end of the course to gauge students' perception of the online online discussions, as well as of the MPEX and the FCI.

77 students participated anonymously in this survey. On a Likert scale, 73% stated that they took the FCI seriously or very seriously, while 65% stated the same about the MPEX. The difference between the answer distributions is however not statistically significant. A larger difference was found regarding the question if the surveys appeared to be relevant: 62% of the students found the FCI relevant, while 51% found the MPEX relevant. These distributions have an α of 1.54, which comes close to confirming a difference at the $p < 0.1$-level.

The most surprising result was that only 31% of the students stated that they would be frustrated or very frustrated if they did not do well on the FCI, and only 30% of the students stated the same for the MPEX. Particularly the FCI percentage is smaller than expected, since the FCI is generally believed to be fairly robust in ungraded settings, see for example Henderson [27], who found only 0.5 points difference between graded and ungraded administration of the FCI. Also, the FCI is similar to the tests and exams used in the course, and students tend to base their relative value system regarding a subject area on the assessments used [15].

On the other hand, student discussions correlate more strongly with performance measures. Students are taking them seriously, likely because they are perceived as helpful and relevant. In the same post-course survey, 89% of the students found the discussions either helpful or very helpful, and 73% stated that they used the discussions to learn physics, as opposed to 35% who said they often or very often just used the discussions to get the correct result as quickly as possible.

While experts would characterize most postings as "bad strategy," only 17% of the students admitted that they often against better knowledge used bad problem solving strategies to get the correct result as soon as possible, and 48% stated that they rarely or never did so (35% were not sure). In the semester after this study was carried out, the author, as a result of the above findings, with a new group of students spent extra time in class demonstrating both good and bad solving strategies, and explaining why bad strategies are bad and eventually work against you. Also, the learning assistants were asked to pay special attention to good problem solving strategies in recitations. At the end of that semester, out of 156 respondents, 3% admitted that they *always* against better knowledge used bad problem solving strategies to get the correct result as soon as possible, 11% stated they often did so, and 52% stated that they rarely or never did so (33% were not sure) — outcomes that are only minimally different from the previous group of students. So, about half of the students claim that their online discussion behavior reflects their epistemological views about physics problem solving; but the other half either "aren't sure" or explicitly admit that their strategies reflect expediency rather than their views about how best to learn physics. These findings correspond to the study by Elby [28], who found that students can perceive "trying to understand physics deeply" and "pursuing good grades" to be a different activities. They also again underline the difference between public and personal epistemologies: the students know that their strategy is bad (public epistemology), but decide it works best for them (personal epistemology), gets results quickly (expediency), and good problem solving behavior does not give them more points as long as they get the correct result (reward structure). Therefore, for the class as a whole, online discussion behavior reflects a combination of students' personal epistemological beliefs, expediency, and expectations about what's rewarded in the class.
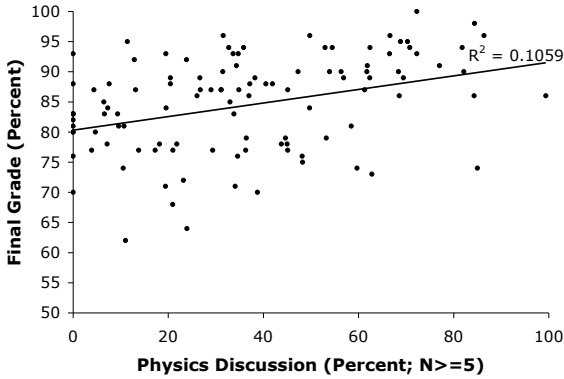
FIG. 2: Correlation of percentage physics-related discussions with grade percentage ($R = 0.33$ $[0.15 \rightarrow 0.49](n = 111)$).
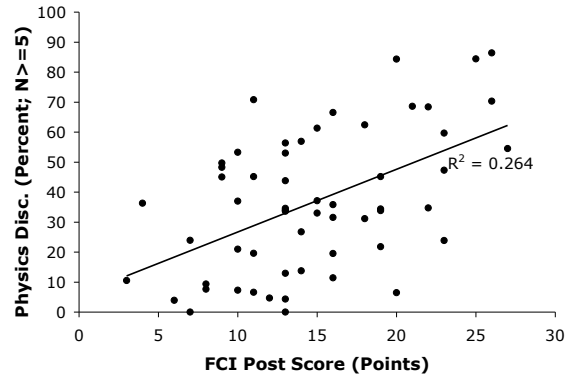

FIG. 3: Correlation of percentage physics-related discussions with final FCI score ($R = 0.51$ $[0.29 \rightarrow 0.68](n = 57)$).

VI. CORRELATION RESULTS

A. Correlations between Discussion Behavior and MPEX

To directly compare the attitudes and beliefs measures, we calculated correlations between the prominence of discussion behavior classes and the MPEX clusters, and generally found them to be very low. As an example, the correlation between the score on the Concepts Cluster and the prominence of conceptual discussion contributions turned out to be $R = 0.14$ $[-0.08 \rightarrow 0.34](n = 84)$ when considering all students, and $R = 0.15$ $[-0.13 \rightarrow 0.41](n = 51)$ when only considering those who made at least five discussion contributions – the 95% confidence intervals (given in square brackets) include zero. Thus, we conclude that discussion behavior and the individual MPEX cluster scores are – if at all – only weakly correlated.

B. Correlations between Discussions and Learning

Figure 2 shows the correlation between the prominence of physics-related discussions and the course grade percentage (for better statistics, only students who contributed at least five discussion entries over the course of the semester were considered).

Figure 3 shows how the percentage of a particular student's discussion contribution that was classified as "physics-related" correlates with their final FCI score ($R = 0.51$ $[0.29 \rightarrow 0.68](n = 57)$). While physics-related discussions positively correlate with FCI scores and grades (Fig. 2), solution-oriented discussions negatively correlate (Fig. 4; $R = -0.58$ $[-0.73 \rightarrow -0.38](n =$
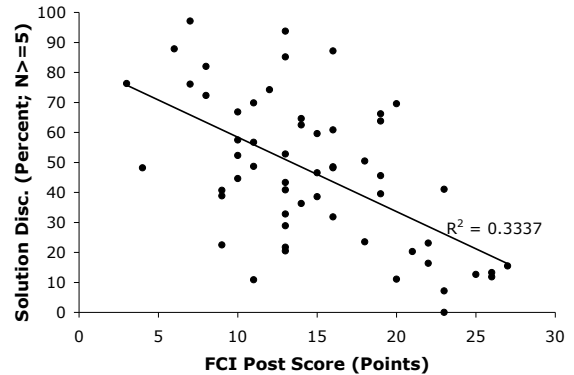

FIG. 4: Correlation of percentage solution-oriented discussions with final FCI score ($R = -0.58$ $[-0.73 \rightarrow -0.38](n = 57)$).

57)).

C. Correlations between MPEX and Learning

Correlations between the MPEX and measures of student learning are generally weak. Considering final exam, FCI, and course grade, $R = 0.36$ $[0.17 \rightarrow 0.52](n = 97)$ between the score on the Coherence cluster and the course grade percentage is the highest correlation found.

Dancy [19] found similarly low correlations with the performance on homework, tests, and final exams: direct comparison with the performance on the final exams found $R = 0.37$ for the correlation with the total
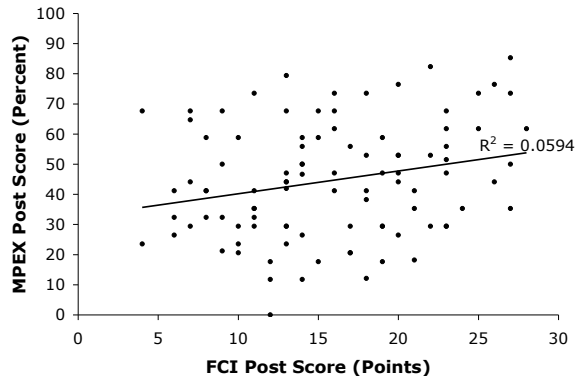
FIG. 5: Correlation of the final FCI score with the MPEX score ($R = 0.24[0.04 \to 0.42](n = 97)$).

MPEX score ($R = 0.27$ here), $R = 0.39$ with the Independence Cluster ($R = 0.25$ here), $R = 0.24$ with the Coherence Cluster ($R = 0.36$ here), $R = 0.29$ with the Concept Cluster ($R = 0.25$ here), $R = -.02$ with the Reality Link cluster ($R = 0.1$ here), $R = 0.3$ with the Math Link cluster (no significant correlation found here), and no significant correlation with the Effort Cluster ($R = 0.1$ here).

Figure 5 shows how the final MPEX and FCI scores correlated with each other, i.e, $R = 0.24 [0.04 \to 0.42](n = 97)$. Coletta and Philips [18] found a strong correlation between the FCI Gain and the MPEX Score ($R = 0.52 [0.24 \to 0.72](n = 37)$), while the same correlation turned out much lower in this study ($R = 0.17 [-0.05 \to 0.37](n = 84)$ here). The correlations reported here are in the same range that Perkins et al. [29] found when investigating the influence of beliefs on conceptual learning, using the CLASS [3] and the Force and Motion Conceptual Evaluation (FMCE) [30] instruments.

VII.   DISCUSSION OF POSSIBLE CAUSAL RELATIONSHIPS

The study showed that there is a relatively strong correlation between solution-oriented discussion behavior (negative) and physics-oriented discussion behavior (positive) and the final FCI score. It is an interesting question whether the students learned physics better because of their more expert-like approach (as argued by Lising and Elby [16]), or vice versa.

In an attempt to answer this question, we are considering the FCI gain as a rough measure of how much physics the students *learned* (versus, for example, knew already). We also introduced a measure of discussion behavior gain

by splitting the semester in half and calculating the the difference between the prominence of discussion behaviors in the first and the second half of the semester.

We then calculated the following two correlations:

- FCI gain versus prominence of solution-oriented and physics-related postings

- FCI gain versus gain in prominence of solution-oriented and physics-related postings

As it turns out, the first correlations are significant, with $R = -0.44 [-0.65 \to -0.18](n = 47)$ for FCI gain versus solution-oriented discussions, and $R = 0.4 [0.13 \to 0.62](n = 47)$ for FCI gain versus physics-related discussions. Such significant correlations do not occur for FCI gain versus any of the MPEX cluster scores.

On the other hand, the correlations with discussion-gain are not significant: $0.24 [-0.05 \to 0.49](n = 47)$ for FCI gain versus gain in solution-oriented discussions, and $-0.12 [-0.39 \to 0.17](n = 47)$ for FCI gain versus gain in physics-related discussions. Note that these correlations have the opposite sign than expected, however, the confidence intervals include zero in both cases. When looking at the absolute values, the average gain in solution-oriented discussions between the two halves of the semester is 2.4%, and the gain in physics-oriented discussions -0.3% — in other words, the students did not really change their discussion behavior over the course of the semester, and their discussion behavior does not improve co-measured with their increasing understanding of physics.

Thus, the discussion behavior appears to be a property of the students that is almost constant over the course of the semester, just like Hammer [5] already pointed out that it is unlikely that epistemological beliefs are changed implicitly by physics instruction.

We also ran a linear regression analysis of the FCI scores versus discussion behavior. In the equations below, "PostFCI" is the predicted post (final) FCI score in points, "PreFCI" is the score on the pre-test FCI in points, and "Solution" and "Physics" are the percentage solution-oriented and physics-related discussion over the course of the semester. For the physics-oriented discussion, we found

$$\text{PostFCI} = 5.486 + 0.922 \cdot \text{PreFCI} + 0.24 \cdot \text{Physics}$$

with an explained variance of 45.6% of the Post FCI score. The effect of the pre-test FCI is significant ($p < 0.001$), the effect of the physics discussion is not ($p = 0.195$).

For the solution-oriented discussion, we found

$$\text{PostFCI} = 7.606 + 0.857 \cdot \text{PreFCI} + (-0.042) \cdot \text{Solution}$$

with an explained variance of 47.9% of the Post FCI score. Both coefficients are significant, the solution-oriented discussion has $p = 0.019$. Thus, controlling for pre-test FCI score, for each 10 percent increase in

solution-oriented discussion, the predicted post-test FCI score goes down by 0.42 points. Students who do not make any solution-oriented contributions would on the average gain 7.6 points on the 30 item FCI due to instruction, while at the other extreme, students who only make solution-oriented discussions would on the average only gain 3.4 points – less than half.

VIII. CONCLUSIONS

Students' online discussion behavior is not a pure reflection of students' epistemological beliefs; other factors, notably expectations and expediency, also feed into their online behavior. This behavior, however, reflects how students actually approach their physics homework problems. Students who exhibit more expert-like approaches have higher learning success, even when controlling for prior physics knowledge. Indeed, the correlation between on-line discussion behavior and conceptual gains is stronger than the correlation between MPEX scores and conceptual gains, showing the value of online discussion behavior as a diagnostic tool.

[1] E. F. Redish, R. N. Steinberg, and J. M. Saul, Am. J. Phys. **66**, 212 (1998).
[2] A. Elby, Am. J. Physics **69**, S54 (2001).
[3] W. Adams, in *Proc. PERC 2004* (2004).
[4] K. Hogan, Science Education **83**, 1 (1999).
[5] D. Hammer, Cognition and Instruction **12**, 151 (1994).
[6] *Using the maryland physics expectations survey*, URL http://www.physics.umd.edu/perg/expects/usempex.htm.
[7] M. T. H. Chi, P. J. Feltovich, and R. Glaser, Cognitive Science **5**, 121 (1981).
[8] A. V. Heuvelen, Am. J. Physics **59**, 891 (1991).
[9] J. Larkin, J. McDermott, D. P. Simon, and H. A. Simon, Science **208**, 1335 (1980).
[10] D. Hammer, Am. J. Phys. **64**, 1316 (1996).
[11] B. K. Hofer and P. R. Pintrich, Review of Educational Research **67**, 88 (1997).
[12] B. K. Hofer, Journal of Educational Psychology Review **13**, 353 (2001).
[13] M. Schommer, Journal of Educational Psychology **85**, 406 (1993).
[14] D. B. May and E. Etkina, Am. J. Phys. **70**, 1249 (2002).
[15] H. Lin, Phys. Teach. **20**, 151 (1982).
[16] L. Lising and A. Elby, Am. J. Phys. **73**, 372 (2005).
[17] A. DiSessa, A. Elby, and D. Hammer, *J's epistemological stance and strategies* (Lawrence Erlbaum Associates, Mahwah, NJ, 2002), pp. 237–290.
[18] V. Coletta and J. Philips, in *AAPT Conference, Summer 2005* (2005), URL http://myweb.lmu.edu/jphillips/PER/AAPT-summer05.pdf.
[19] M. Dancy, in *Poster, PER Conference 2002* (2002), URL http://www.physics.uncc.edu/physstaff/mhdancy/papers/poster2002_reasoning.ppt.
[20] G. Kortemeyer, M. Hall, J. Parker, B. Minaei-Bidgoli, G. A. II, W. Bauer, and E. Kashy, Journal of Asynchronous Learning Networks **9** (2005).
[21] G. Kortemeyer, Am. J. Phys. **74**, 526 (2005).
[22] D. A. Kashy, G. Albertelli, W. Bauer, E. Kashy, and M. Thoennessen, Journal of Asynchronous Learning Networks **7**, 31 (2003).
[23] G. M. Novak, E. T. Patterson, A. D. Gavrin, and W. Christian, *Just-in-time teaching: blending active learning with web technology* (Prentice Hall, 1999), ISBN 0-13-085034-9.
[24] *The learningonline network with capa*, URL http://www.lon-capa.org/.
[25] D. A. Kashy, G. Albertelli, G. Ashkenazi, E. Kashy, H.-K. Ng, and M. Thoennessen, in *Proc. Frontiers in Education* (2001), vol. 31, pp. 1073–1078, ISBN 0-7803-6669-7.
[26] I. Halloun, R. R. Hake, E. P. Mosca, and D. Hestenes, *Force concept inventory*, URL http://modeling.la.asu.edu/R&E/Research.html.
[27] C. Henderson, Phys. Teach. **40**, 542 (2002).
[28] A. Elby, Am. J. Phys. **67**, S52 (1999).
[29] K. Perkins, , W. Adams, N. Finkelstein, S. Pollock, and C. Wieman, in *Proc. PERC 2004* (2004).
[30] R. Thornton and D. Sokoloff, Am. J. Phys. **66**, 338 (2005).
[31] For example, a student evaluation of the course under consideration in this study stated: "Don't derive anything, really, you have a Ph.D. I'll believe what you tell me."