Work in Progress: PEERSIST—An Observational Study of Student Questions to Identify Levels of Cognitive Processing

Sarah Johnston, Arizona State University
Cody D. Jenkins, Arizona State University
Ms. Thien Ngoc Y. Ta, Arizona State University, Polytechnic Campus

Thien Ta is a doctoral student of Engineering Education Systems and Design at Arizona State University. She obtained her B.S., and M.S. in Mechanical Engineering. She has taught for Cao Thang technical college for seven years in Vietnam. She is currently

Dr. Ryan James Milcarek, Arizona State University

Ryan Milcarek obtained his B.S., M.S. and Ph.D. in the Mechanical & Aerospace Engineering Department at Syracuse University. He also obtained a M.S. in Energy Systems Engineering, Certificate of Advanced Study in Sustainable Enterprise and Certificate in

Dr. Gary Lichtenstein, Arizona State University

Gary Lichtenstein, Ed.D., is founder and principal of Quality Evaluation Designs, a firm specializing in education research and program evaluation.

Dr. Samantha Ruth Brunhaver, Arizona State University, Polytechnic Campus

Samantha Brunhaver is an Assistant Professor of Engineering in the Fulton Schools of Engineering Polytechnic School. Dr. Brunhaver recently joined Arizona State after completing her M.S. and Ph.D. in Mechanical Engineering at Stanford University. She also

Dr. Karl A. Smith, University of Minnesota, Twin Cities

Emeritus Professor of Civil, Environmental, and Geo- Engineering, Morse-Alumni Distinguished University Teaching Professor, Faculty Member, Technological Leadership Institute at the University of Minnesota; and Emeritus Cooperative Learning Professor of Engineering Education at Purdue University..

Work in Progress: PLSG – An Observational Study of Student Questions to Identify Levels of Cognitive Processing Reached During Discussion

Abstract - This Work in Progress study reviews the PEERSIST (PEER-led, Student Instructed, STudy group) project, which explores the achievement and persistence of students who experience the Peer-Led Study Group (PLSG) model being implemented in an undergraduate thermodynamics course at Arizona State University (ASU). In this study, we analyze students' levels of cognitive processing reached during weekly recitations, in which they are grouped in PLSGs. Specifically, we explored whether the proposed observational protocol regarding students' question-prompted discussion captures multiple levels of cognitive processing. The questions asked by students and resulting group discussions were categorized based on the levels of the cognitive processing dimension of Bloom's revised taxonomy of learning. PLSG groups were identified based on within-group demographics (i.e., white men, women, international, and Hispanic/Latino students), with the same group being observed for multiple weeks. Session excerpts were recorded, and exchanges were coded using Bloom's revised taxonomy.

Keywords - Peer instruction, Bloom's revised taxonomy, Cooperative learning, Study groups, Observation protocol

INTRODUCTION AND BACKGROUND

The PLSG (Peer-Led, Study Group) model has been implemented at Arizona State University (ASU) both in part and entirely within an undergraduate thermodynamics course, along with parallel research regarding student pass rates and other factors [1]. In the PLSG model, students work together in small groups of four to five peers to solve challenging, course-related problems during weekly, 50-minute recitation sessions. While a facilitator observes and offers assistance to keep the group on track, the PLSG model relies upon student-student discussion and interactions as the primary learning method instead of direct instruction from an instructor or a TA. Implementing the PLSG model has improved students' course grades and pass rates compared to the typical TA-led Recitation (TAR) method, which involves approximately 30 students watching a TA solve problems on a whiteboard [1].

The PLSG model is based on Treisman's work at UC Berkeley [2]. He observed Asian students working in self-formed study groups, wrestling through problems in a gateway calculus course, and hypothesized that by working together, students were co-constructing disciplinary knowledge so that they began to think like mathematicians. He instituted similar study groups for African-American students, which turned the tide on their high failure rates. Treisman's model has been implemented in universities nationwide since, with consistently powerful effects, including at the University of Texas, Austin, where he currently teaches.

Despite the demonstrated success of PLSGs over the past 40 years, we have yet to find empirical evidence that the model's effectiveness has resulted from peer interactions. The current study sought to capture peer discussion features reflective of discipline-based cognitive processing. We hypothesized that when group members asked questions and had discussions at higher levels of the cognitive processing dimension of Bloom's revised taxonomy, a tool for categorizing these levels, they engaged in higher-order cognitive processing [3]. Thus, we ask in this study, "In

what ways can a real-time, repeated-measures observational protocol based on Bloom's revised taxonomy effectively capture and categorize distinct levels of cognitive processing?"

In assessing student subject cognition, there are a few main types, with examples including "diagnostic, formative, and summative [4]." Diagnostic assessment determines students' current knowledge, formative assessment occurs through feedback during the educational process, and summative assessment occurs after learning, often through reviews concluding a course [4]. Our research uses formative assessment to improve future and current course semesters. With the information collected from recording students' questions, we hope to find a connection between the types of questions they ask and their achievement based on pre-course preparation to identify potential reasons for the disparity in their course pass rates.

Because our literature review did not reveal any standardized methods of reviewing and categorizing questions asked by students, we developed our own protocol based on the categories of the cognitive processing dimension of Bloom's revised taxonomy [3,5]. This paper describes the observation protocol and methodology for assigning students' question-prompted discussions and resulting discussions to different levels of cognitive processing.

METHODOLOGY

For the thermodynamics recitations, students were first grouped based on their beginning-of-term (BoT) GPA. Then, white men, women, international, and Hispanic/Latino students were grouped together into separate groups. This study involved 23 undergraduate students comprising six of the mentioned groups. Of these, the first four groups had a minimum average GPA of 3.9 and consisted of 1) four white men, 2) four women, 3) three international and two unknown minority status students, 4) one Hispanic/Latino student and two white students. These groups are referred to as the 4.0 students for this study. The other two groups included 1) two Hispanic/Latino students and one student of two or more races with an average GPA of 3.46 and 2) three Hispanic/Latino students and one white student with an average GPA of 2.65.

When selecting which groups to observe, we grouped women and men separately to determine whether gender affected what types of questions were being asked within recitations. International students were grouped together because best practice has shown that grouping students with similar demographics together, such as multiple women in a group, can promote participation [6, 7]. Furthermore, approximately 11% of the students in the course are international students, a similar proportion compared to all engineering students at ASU [8]. Finally, because ASU is a designated Hispanic Serving Institution (HSI), Hispanic/Latino students were selected and compared across different incoming GPAs to assess the impact of pre-course preparation on student cognition during study groups compared to other demographics.

After assigning students to groups, a series of 21 data collection periods were conducted. Each data collection period consisted of a 20-minute observation, during which all questions students asked were recorded for later analysis. Observations occurred at different times during weekly 50-minute recitation periods, all taking place on the same day of the week. During observation, the questions students asked and the discussion following each question were recorded. This context was important in helping us categorize students' questions based on the perceived

cognitive level of learning reached within the prompted discussion. These levels of cognition were based on the cognitive process dimension of Bloom's revised taxonomy [3,5]. The skills needed to reach different levels of cognitive processing are included in Table 1 below, with *remember* representing the most basic level and *create* representing the most advanced [3]. Although not included in this paper, Bloom's revised taxonomy also features a knowledge dimension, which organizes levels of information to be demonstrated, including *metacognitive*, *procedural*, *cognitive*, and *factual* [3].

Table 1. Bloom's revised taxonomy for cognitive processing based on [5]

Level	Associated Verbs
Create	Design, assemble, construct, conjecture, develop
Evaluate	Appraise, argue, defend, judge, critique
Analyze	Differentiate, organize, compare, contrast, question
Apply	Execute, solve, use, interpret, demonstrate
Understand	Discuss, explain, recognize, classify, identify
Remember	Define, state, repeat, memorize, duplicate

Recorded questions and resulting discussions were matched with associated verbs from the different levels of Bloom's revised taxonomy for cognitive processing. If verbs from multiple categories fit a question, the category with the most applicable verbs was assigned following a two-party discussion regarding categorization. Along with the categorizations based on Bloom's revised taxonomy, the team developed three additional levels of question organization: *social* questions, *logistic* questions regarding problem-solving, and questions for *checking the value* of a variable. From our analysis, questions and their associated discussions reached as high as the *evaluate* level. Examples of questions from each observed category are shown in Table 2.

Table 2. Example questions for each category

Category	Question	Associated Verb(s)
Social	How do you guys feel about the exam next week?	NA
Logistic	Does someone want to take over [solving the problem]?	NA
Checking the Value	Is that what you got?	NA
Remember	What's the equation?	State
Understand	So would we do the energy equation to find the Q [heat transfer]?	Recognize
Apply	How did you convert to kilowatts?	Solve, Sketch
Analyze	Why can't we use enthalpy?	Differentiate
Evaluate	Wait, how do we have an adiabatic turbine?	Argue, Defend

Categorizing questions depends on more than the initial question. A question like, "How did you convert to kilowatts?" may not appear to be an *application* level of learning initially. Following the question, the study group went through the full unit conversion together, deriving the units from first principles, to explain the concept to their team member(s). They fully solved the conversion and wrote it out to ensure everyone understood. Because the question led to using images (sketch) and applying a formula to convert a result into kilowatts (apply), the event was coded as *apply*. Including the discussion following a question in the analysis enabled us to more

accurately capture the cognition level associated with a question and also allowed us to review correlations between beginning-of-term (BoT) GPA and the average number of questions asked of each type on both individual and group levels.

RESULTS

Figure 1 below includes the average number of question-prompted discussions per observation compared to the individual students' beginning-of-term GPA. Cognitive levels above *understand* were rarely recorded. However, this result was not surprising as the course learning objectives only required students to reach the *understand* level of cognition. There are higher concentrations of data around the 4.0 GPA range because more groups were selected at this range originally. With further data collection, we hope to see greater concentrations of *apply*, *analyze*, and *evaluate* question-prompted discussions, as well as data across a wider range of beginning-of-term GPAs. Nonetheless, the figure below concludes that the proposed method was useful in tracking all eight levels of cognitive processing listed in Table 2.

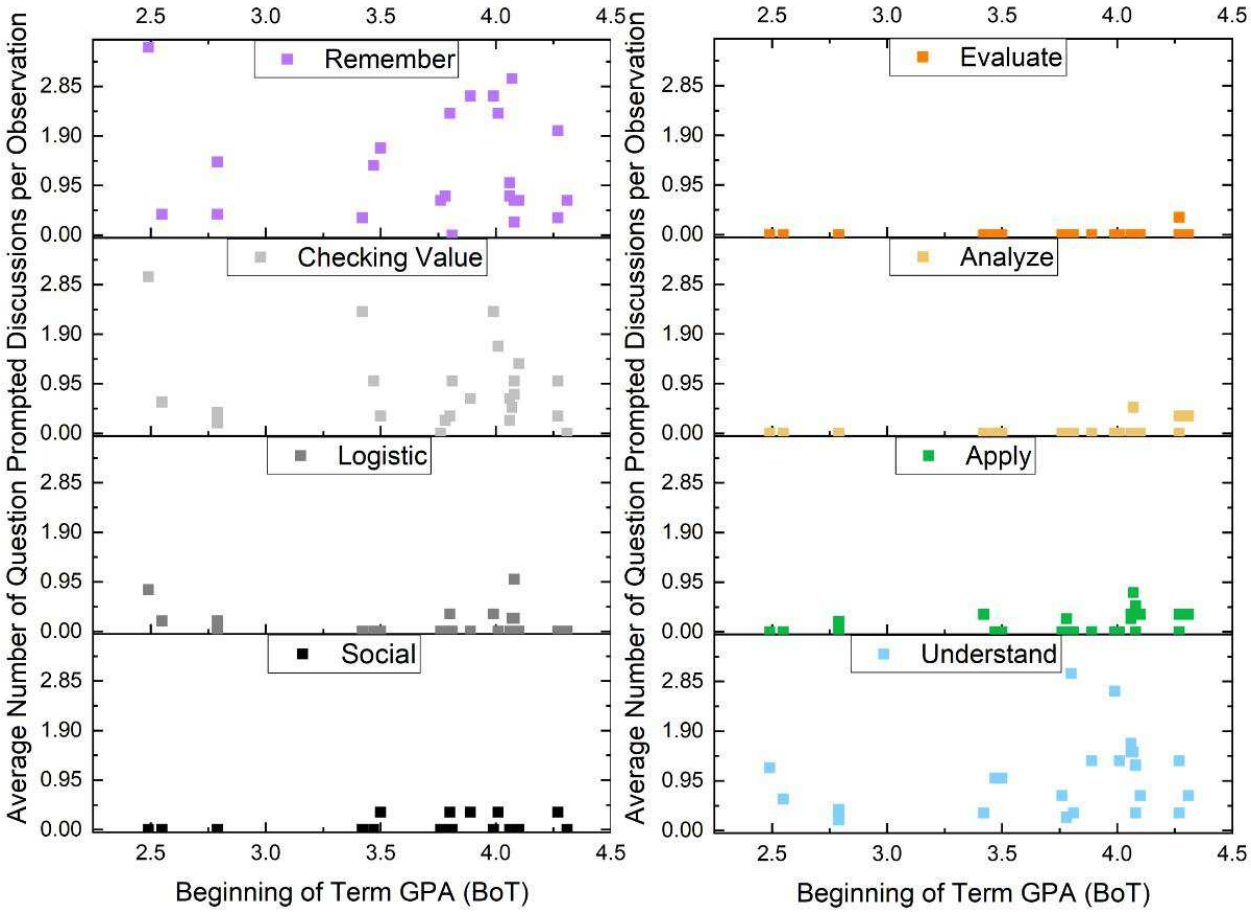

Figure 1. Average number of question-prompted discussions per observation by GPA

With confidence in the observation protocol and methodology, we began analysis to answer our second research question, "To what extent do student pre-course preparation and student demographics influence the questions asked and resulting discussion?" Figure 2 shows this data for the 4.0 GPA groups only. Despite these groups having different demographics, we saw

similar distributions of question-prompted discussion types in the *remember* and *understand* categories and in the total number of questions asked per observation. We see this as a positive result because each group had similar levels of cognitive processing when beginning-of-term GPA was held constant, regardless of other demographic considerations.

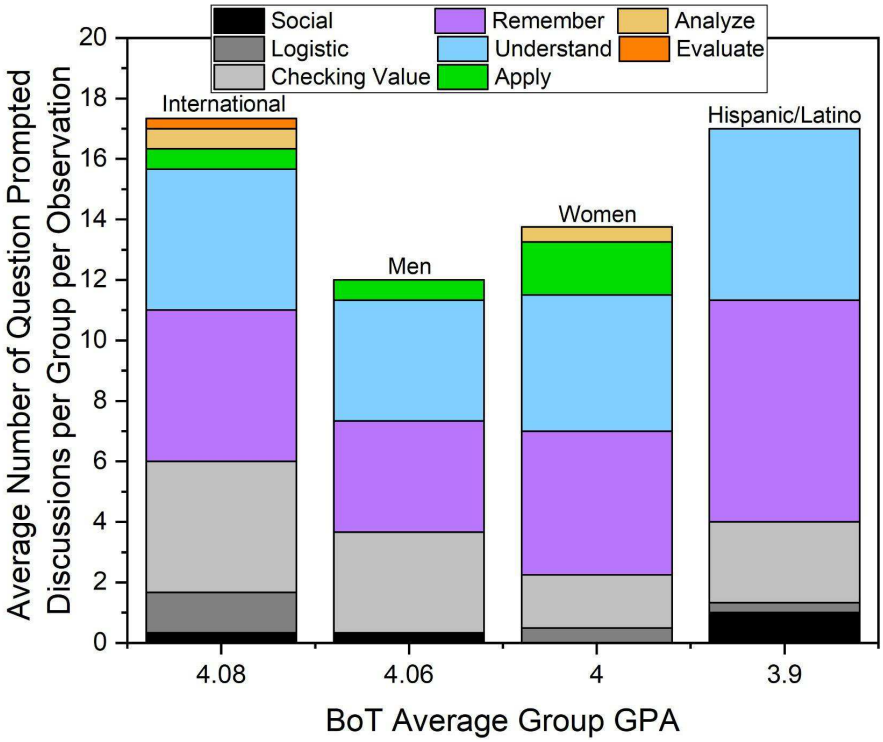

Figure 2. Average number of question-prompted discussions per group per observation vs. average beginning-of-term group GPA for 4.0 GPA groups only

Data were insufficient to fully evaluate the impact of pre-course preparation on levels of cognitive processing during PLSGs for lower GPA student groups. However, a previous observation study of student interactions in PLSGs showed that all students had a similar frequency of interactions, regardless of GPA [9].

DISCUSSION & FUTURE WORK

Based on course outcomes, we hypothesized that students could reach as high as the *understand* level within recitation, but these expectations were surpassed by some students, reaching as high as the *evaluate* level. These positive results validated our hypothesis that the developed protocol could accurately track multiple levels of cognitive processing during a PLSG and indicate that we can further our data collection to answer 1) "To what extent do student pre-course preparation and student demographics influence the questions asked and resulting discussion?", and 2) "What level of discipline-based cognition are students reaching based on questions asked and discussion occurring within the PEERSIST model?" We are addressing these questions by employing the protocol with a larger sample of PLSG groups conducted over an entire semester. In addition, the knowledge dimension of Bloom's revised taxonomy will be analyzed to study if students are learning in a *metacognitive*, *procedural*, *conceptual*, or *factual* manner.

ACKNOWLEDGEMENT

This material is also based upon work supported by the National Science Foundation under Grant No. 2205001.

REFERENCES

[1] R. J. Milcarek, S. R. Brunhaver, T. Ta, C. Jenkins, G. Lichtenstein, and K. A. Smith, "The Impact of Peer-led Study Groups on Student Achievement in a Gateway Engineering Thermodynamics Course." available by request.

[2] E. Hsu, T. J. Murphy, and U. Treisman, "Supporting High Achievement in Introductory Mathematics Courses: What We Have Learned from 30 Years of the Emerging Scholars Program," in Making the Connection, The Mathematical Association of America, 2008, pp. 205–220. doi: 10.5948/UPO9780883859759.017.

[3] L. Anderson and D. Krathwohl, Eds., *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*, NY, USA: Longman, 2001

[4] Northern Illinois University Center for Innovative Teaching and Learning, "Formative and Summative Assessment." https://www.niu.edu/citl/resources/guides/instructional-guide/formative-and-summative-assessment.shtml (accessed Dec. 15, 2024)

[5] P. Armstrong, "Bloom's Taxonomy." https://cft.vanderbilt.edu/guides-sub-pages/blooms-taxonomy/ (accessed Oct. 10, 2023)

[6] S. V. Rosser, "Group Work in Science, Engineering, and Mathematics: Consequences of Ignoring Gender and Race," *College Teaching*, vol. 46, no. 3, pp. 82-88, Aug. 1998, doi: http://www.jstor.org/stable/27558891

[7] A. A. Niler, R. Asencio, and L. A. DeChurch, "Solidarity in STEM: How Gender Composition Affects Women's Experience in Work Teams," *Sex Roles*, Apr. 2019, doi: https://doi.org/10.1007/s11199-019-01046-8

[8] Ira A. Fulton Schools of Engineering, " Enrollment and Degrees Granted." https://engineering.asu.edu/enrollment/ (accessed Dec. 15, 2024)

[9] C. Jenkins, T. Ta, R. J. Milcarek, G. Lichtenstein, S. R. Brunhaver, and K. A. Smith, "Work in Progress: An observational study of the PEER-led, Student Instructed, STudy group (PEERSIST) model in thermodynamics."