AN ABSTRACT OF THE THESIS OF

Jessie Keeler for the degree of Master of Science in Chemical Engineering presented on May 11, 2016.

Title: Reflection and Dissection in Engineering Education: Exploring Critical Elements of Active Learning and their Roles in Course Structures

Abstract approved:

Milo D. Koretsky

This thesis focuses on active learning through four different studies that include metacognitive and cognitive aspects of learning designs to support active learning and a contextual analysis of the implementation of teaching tools used in active learning. The first three investigate important elements of active learning whereas the fourth study explores how those elements fit together. The first two studies evaluate how student responses vary based on different reflection prompts. The third study examines student thinking processes as they work through interactive simulations. Through a comparative case study, the fourth study analyzes how two different instructors implement similar active teaching tools into their respective courses.

Reflection and Dissection in Engineering Education: Exploring Critical Elements of Active Learning and their Roles in Course Structures

by
Jessie Keeler

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented May 11, 2016
Commencement June 2016

Master of Science thesis of Jessie Keeler presented on May 11, 2016

APPROVED:

Major Professor, representing Chemical Engineering

Head of the School of Chemical, Biological, and Environmental Engineering

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Jessie Keeler, Author

ACKNOWLEDGEMENTS

Thank you to my family and friends for reminding me what’s important in my life, to everyone who contributed to this writing, and to the animals in my life for keep me sane.

CONTRIBUTION OF AUTHORS

For the study described in Chapter 2, Bill Brooks performed statistical analysis on the results and contributed to writing the Results and Discussion sections of the manuscript. Debra Gilbuena contributed writing to the Introduction, Background, and Discussion sections. Jeffery Nason contributed his perspective as the instructor in which the reflection activity was implemented within the Results section of the manuscript.

For the study described in Chapter 4, Ying Cao contributed to the data analysis, framing, and overall writing of the manuscript. Thomas Ekstedt developed the referenced grading system within Concept Warehouse and generated Figures 4.6 and 4.7 for use within the manuscript.

For the study described in Chapter 5, John Ivanovitch performed the data collection and analysis for the biology course; he also contributed to the overall writing throughout the manuscript. Jana Bouwma-Gearhart will contribute to the Conclusion section of the manuscript.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF FIGURES

LIST OF TABLES

1. GENERAL INTRODUCTION

Active learning has been gaining attention because numerous studies show it increases both student learning gains and student retention.[1,2,5] The goal of these learning techniques is to improve student learning by actively engaging them through direct participation in the learning process as opposed to passively listening to lectures. Extensive research on active learning has shown it helps improve education, but the wide range of contexts it is studied in also makes defining active learning difficult. For clarification, within this paper "active learning" will be defined using Prince's[3] definition, which is "any instructional method that engages students in the learning process." This definition can be expanded by acknowledging there are different levels and types of engagement. For example Chi and colleagues[4] propose the ICAP framework where student engagement can be defined as passive, active, constructive, or interactive based on the task that student is performing. According to Chi[5], the more engaged a student is the more they learn and the better they understand the material. Therefore, active learning requires students to be engaged in the learning process, but not all active learning methods will prompt the same level of engagement.

As explicitly stated in some of the subsequent chapters, the goal of the studies in this thesis is not to determine whether or not active learning, or more specifically active learning, is beneficial for students. Many studies have already proven how valuable engagement is for student learning. For example, Hake[1] compared student performance data between students in physics courses that implemented active learning, or interactive engagement, and students enrolled in traditional physics courses. Overall, Hake discovered that the gains between pretest and posttest scores for students in courses that implemented active learning were double the gains of students in traditional courses.[1] Similarly, Freeman et al.[2] discovered through a meta-analysis of 642 papers that students in a traditional course are 1.5 times more likely to fail than students in an equivalent course that implements active learning. These findings hold true across all STEM disciplines, regardless of level or type of the course.[2]

The four studies presented in this thesis aim to examine different elements involved in active learning and what kind of thinking processes and responses they elicit from students, or, in the case of one study in particular, what kind of thinking processes instructors intend for these elements to elicit from students. Together, these four studies paint a more nuanced picture of what active learning looks like in practice as well as a detailed understanding of how different student-centered activities require students to be engaged.

Chapters 2 and 3 describe a set of studies focused on a specific active learning technique, a reflection activity, by comparing different ways it can be implemented. Promoting students to be more reflective about their learning experiences allows them to develop robust learning strategies and develops the metacognitive skills that are a characteristic of expertise. It can also provide instructors formative information about students conceptions of course content. We use the definition that reflection is the act of "exploring the meaning of experiences and the consequences of the meanings for future action." A simple reflection activity commonly used in class is a Muddiest Point exit question, "What was the muddiest point in class this week?" In this activity, the instructor asks students to write a brief, anonymous written comment describing the concept or topic that they found to be the most difficult to understand during class.

In the study described in Chapter 2, we utilized two different reflections prompts: the classic "Muddiest Point" prompt as well as a "Most Surprised" prompt. We then analyzed how student responses varied based on which prompt they responded to. Through our examination of student responses, we concluded that the different reflection prompts elicited different types of thinking about the class.

Based on the analysis of student reflection responses in that study, we developed and implemented a hybrid reflection activity that allowed students to choose amongst a "Muddiest Point" prompt, a "Most Surprised" prompt, or to use both. In Chapter 3, we describe a similar analysis of student responses and conclude that each prompt elicits different responses and provides unique benefits.

Chapter 4 describes a study in which we observed how students think through different frames in interactive simulations known as an *Interactive Virtual Laboratories* (IVL). We audio recorded students completing the simulation and then analyzed the recordings along with the students' submitted answers. This allowed us to understand the thinking processes that led students their submitted answers. We were able to identify students who were not audio recorded but likely had similar misconceptions by comparing their numerical answers to those of the students who were recorded. These analyses allowed us to better understand common misconceptions in the topic as well as determine approximately how many students shared those misconceptions. This study also aimed to develop an automatic grading system for the IVLs and use the data gathered from student responses to better inform instructors about where students are struggling.

Chapter 5 details a study in which we examined how instructors in two different courses implemented similar active learning tools. Specifically, we focused on what kind of thinking processes they intended these tools to elicit from students. Through the analysis of course artifacts and instructor interviews, we sought to provide a thorough description of what constitutes active learning and some of the ways it can be incorporated into the classroom learning environment.

2. WHAT'S MUDDY VS. WHAT'S SURPRISING? COMPARING STUDENT REFLECTIONS ABOUT CLASS

Jessie Keeler, Bill J. Brooks, Debra M Gilbuena, Jeffery A. Nason, & Milo Koretsky

Proceedings of the 2015 American Society for Engineering Education Annual Conference & Exposition
https://peer.asee.org/what-s-muddy-vs-what-s-surprising-comparing-student-reflections-about-class

Introduction

Classroom assessment techniques that ask students to reflect on material covered in class are believed to help improve learning by allowing the student to actively participate in the learning process while evaluating their understanding of course content.[1] Promoting students to be more reflective about their learning experiences allows them to develop robust learning strategies and metacognitive skills that are characteristic of expertise.[2] Students' written reflections can also provide instructors with formative information about students conceptions of course content. In this paper, we use the definition of Turns and Atman that reflection is the act of "exploring the meaning of experiences and the consequences of the meanings for future action."[3]

Instructors have used several forms of short reflection activities at the end of class. In this paper we compare two such activities which we term *Muddiest Point* and *Most Surprised*. Muddiest Point asks the students to answer the exit question, "What was the muddiest point in class this week?" In this activity, the instructor asks students to write a brief, anonymous written comment describing the concept or topic that they found to be the most difficult to understand during class. Similarly, in the Most Surprised reflection activity, the instructor asks students to answer the question "What surprised you most about class this week?" In this study, we use a quasi-experimental design to empirically investigate students' in-class responses to these two end of in-class activities where students were asked to reflect on the class over the last week. One recitation section is provided the more common *Muddiest Point* exit question. The alternate section is provided the *Most Surprised* exit question. We ask the research question, “How do the student reflection responses differ based on the type of exit question asked?”

Our hypothesis is the different reflection prompts will elicit different types of thinking about the class, but that both will provide the learners and the instructor productive information. By understanding the differences in how students answer these reflective exit questions, instructors can more intentionally select exit questions appropriately.

Background

Instructors have used several short end-of-class reflection activities to both promote student learning and gather formative information about students’ conceptions of course content. Prompts include requesting students to describe the most important,[4-6] interesting,[7] confusing or muddiest,[5-10] and surprising[11] aspects from a lecture, recitation, or week of the course. In this paper we focus on the latter two.

Some of the early uses of the Muddiest Point activity focused on providing formative information for the instructor. For example, Mosteller[4] reported on the use of the Muddiest Point activity in a statistics class and argued that the practice “…promotes concrete and sometimes non-trivial responses to the question of what ‘you want to know more about’”(p. 19). A few years later Angelo and Ross[5] described the muddiest point as having one of the best returns on investment; although very little time is required to perform a Muddiest Point implementation, it can yield useful information for evaluating a class.

More recent studies emphasize the use of the Muddiest Point activity to promote student learning, particularly metacognition. For example, Tanner[7] found that assigning a Muddiest Point activity helps students realize that being confused is a natural aspect of learning. This activity also was found to allow the students to safely share what they find unclear without having to reveal their confusions to the entire class and thus risk judgment. Tanner argued that while students might be unaccustomed to being open with their professors about their confusions, the process of acknowledging, embracing, and resolving their misconceptions is a key to evolving as learners.

Hall and colleagues[6] report on a study in the context of increasing active learning in an engineering class. That study incorporated activities such as concept tests, group discussions, and Muddiest Point evaluations across several different engineering disciplines. They found that a majority of professors and students found Muddiest Point evaluations to be effective. Students felt that the professors' clarifications about the Muddiest Points directly improved their learning, showed their professors cared, and enhanced their overall relationship with their professors. The exercise was also popular amongst professors; several planned to include the exercise in their future courses. In Introductory Materials courses, Krause and colleagues[8,9] found the use of Muddiest Point activities informed instructors' use of formative process feedback and improved student attitudes, achievement and retention of course content.

Most Surprised activities are rarely used in engineering, but have been used by instructors in other fields in a way similar to Muddiest Point. In the most common form reported in the literature, Most Surprised is posed as one question in Brookfield's[11] five-question critical incident questionnaire (CIQ). The CIQ is intended to be assigned to students at the end of a course period or week of a course, in a similar manner to the Muddiest Point activity. In the CIQ, Most Surprised comes as the fifth and final item and is preceded by questions regarding when students felt most engaged and least engaged, and what actions taken in class were most confusing and most affirming.

The use of the CIQ has been reported in many different educational settings, including health education,[12] adult online education,[13] public communication,[14], writing,[15] management education,[16] and engineering mechanics.[17] Generally the CIQ has been described as a useful instructional tool that promotes learner reflection. However, studies that explicitly assess the tool's effectiveness in various settings are lacking.[18]

Hessler and Taggart[15] reported that the CIQ solicited responses related to issues with pedagogical approaches rather than related to course content. In addition, they suggest that students' regular completion of the CIQ, including the Most Surprised question, helped students develop habits of a reflective practitioner and gave the instructor information to improve instruction. They noted that responses to the Most Surprised question were often related to times students reported feeling most engaged, disengaged, or affirmed in the course. While most implementations have used the original questions, Keefer[19] suggested a revision of the CIQ that omits the Most Surprised question and instead asks students to identify the most important information learned in class and solicits questions or suggestions from the instructor. However, this revision has not been well tested.

While both the Muddiest Point and Most Surprised have been noted as providing instructors with formative assessment information and promoting students' critical reflection, there is little research on using Most Surprised as a sole reflection prompt and also little research comparing these exit questions. Our study seeks to provide a better understanding of the types of responses elicited by these questions so that instructors can select the exit question appropriately.

Methods

Our quasi-experimental study empirically investigates students' in-class responses to weekly reflection prompts. Participants were enrolled in a sophomore-level course titled "Material Balances and Stoichiometry" at a large public university in the Northwest United States. The course is a requirement of chemical engineering, bioengineering, and environmental engineering degree programs. It is the first of a three-course sequence that is followed by "Energy Balances" and "Process Data Analysis." Being the first department-specific requirement in the curriculum, the course also serves as the entry point for transfer students. Data are only reported for students who agreed to participate and signed an informed consent form approved by the Institutional Review Board.

The students attended a common lecture and self-selected into one of two weekly, one-hour recitations. There were 117 and 150 students that consented and wrote at least once response in recitation sections 1 and 2, respectively. Each week students were asked to provide responses to one of the following reflections in class:

1. What was the muddiest point in class today/this week? (Muddiest Point)
2. Describe what surprised you most in class today/this week (Most Surprised)

The questions were posed to students in the last five minutes of recitation with the objective that students would reflect on all of the previous week's activities. Each recitation section was asked to respond to both prompts, but in alternating weeks. So in a given week, one recitation section would answer the Muddiest Point while the other answered the Most Surprised. For example in Week 2, students in the recitation section 1 were asked the Muddiest Point and section 2 the Most Surprised. During the subsequent weeks, the Muddiest Point and Most Surprised were assigned alternatingly to the recitation sections. This research design allows us to compare the resulting student reflections based on the same content and coverage. Students provided responses on their laptops, smartphones, or tablets using the *Concept Warehouse*[20] where they were stored in a database and available for analysis.

Word count data and thematic codes and were obtained from the collected responses. The word count data served as a proxy measure of engagement. These data were used to determine if one of the questions prompted more elaboration from students than the other question. The thematic coding provided a qualitative analysis of the student responses. This assessment allowed us to compare the content of the responses each prompt elicited. As stated above, Most Surprised in the context of the CIQ seldom provided content specific elaboration. We are interested to see the proportion of responses that address content relative to structural and pedagogical issues and, of those that address content, how they compare to the Muddiest Point. The coding process also provided information relative to affective differences in the responses.

Word counts for each of the approximately 1600 responses were computed and used for broad comparison of the entire dataset. We then used an emergent coding scheme on the subset from second and fifth weeks to more finely compare the results elicited by the two different prompts. These weeks were chosen because we had complete data sets (i.e., each section answered one of each reflection question), they were not the week of an exam, and they were as early in the term as possible.

We analyzed responses using open coding, a process used to infer categories of meaning. The coding process involved reviewing the responses and sorting them into categories. Initially, each set of data was scanned for reoccurring responses that could be used to generate the coding categories. After the first set of categories were created, there were a few responses left without a category. If these responses could be related to one another, a new category was created for them. When all of the remaining responses were unconnected both to each other and to existing categories, they were coded with the catch-all category “Other.”

Categories that were general and applied to both Week 2 and Week 5 data are shown in Table 2.1, along with a description of the code and examples reflections from each prompt. Not all codes were observed for both prompts; in such a case, only one example is given. Categories for week specific content are shown for Week 2 and Week 5 in Tables 2.2 and 2.3, respectively. For example, during the second week the concept of buoyancy was frequently identified in both the Muddiest Point and Most Surprised. The categories in Table 2.1 are general and can be interpreted in the context of any engineering science course. The categories in Tables 2.2 and 2.3 are topic-specific and meant to provide a comparative example of how content changes with prompt.

The first 50 responses for each prompt (100 total) in Week 2 were independently coded by two of the authors. Inter-rater reliability using the Cohen’s Kappa statistic is 0.94 for Muddiest Point and 0.92 for Most Surprised. This shows acceptable reliability for the coding process.

Table 2.1 General coding categories for reflection prompts.

Category	Description	Examples	
		Most Surprised	Muddiest Point
Class/HW	Responses relating to the structure of the class, studios, or homework that has been assigned.	How fast you went in Monday's lecture. By the end I was pretty much writing down symbols I didn't know the meaning of.	I think the lectures were overall good, but there are some problems in the hw that was hard
Misconceptions	Responses in which students explained the material introduced in the class changed the way they previously thought about certain concepts/ideas.	I'm most surprised by the fact that for the conceptual questions the answer sometimes goes against my first instinct.	
Nomenclature	Responses involving the symbols or terms used in the class. This category also includes trouble writing symbols correctly.	I was surprised by the amount of vocabulary that goes into defining systems and processes	The muddiest point was pounds/pound mole. Is pound mole a completely different thing than pound*mole? It's just a strange unit.
Previous Knowledge	Responses relating the material in class to concepts that a student has learned in a previous class.	I really enjoy the chemistry concepts. I was surprised most about how you cane easily relate chemistry into the systems	
Problem Solving	Responses relating to the process of solving problems within the class.	That balancing processes is a much simpler process when looked at one step at a time.	setting up the equation to answer for the flow rate for problem C
Test	Responses focusing mainly on midterm exams.	That we have a midterm next week.	My midterm score ....
Specific Content	Responses relating to specific content covered in class. See Tables 2.2 (Week 2) and 3 (Week 5).	See Tables 2.2 and 2.3	See Tables 2.2 and 2.3
Positive	Responses with positive connotations.	How I'm starting to finally understand more!	At first the weird extent of reaction, but now I think I have a firm grip on that!
Negative	Responses with negative connotations.	most of the lecture material is easy to understand but the work part seems much harder and this freak me out.	I thought that studio was quite difficult - I would appreciate a hint or a general overview of concepts we will need before we begin.
Neutral	Responses with neither positive nor negative connotations.	that when ice met it does not change the over all water level of a glass afterwards.	Differentiating between the types of process'
Other	Responses that have nothing to do with the class or the material that was covered.	I was surprised by the amount of people talking during our lecture.	When it rained outside, that was pretty muddy.
Nothing	When a student finds nothing surprising or muddy.	Nothing is really surprising in this week.	I understood everything fine

Table 2.2 Week 2 content specific coding categories for reflection prompts

Week 2 (Specific Content)			
Category	Description	Examples	
		Most Surprised	Muddiest Point
Balances	Responses focusing on the process of performing material balances on a system	The in-depth process that needs to be taken to figure out the material balance equation	
Buoyancy	Responses involving buoyancy and other related principles such as Archimedes' principle and buoyant forces	Throwing the rock out the boat and having the water displaced decrease instead of staying the same or increasing.	The problems that involve buoyant force like the one with melted ice.
Mole/Mass	Responses focusing on concepts of mole/mass fraction		I have some difficulty converting from mass fraction to mole fraction etc...
Pressure	Responses regarding pressure concepts.	Differences in the types of pressure(gauge, absolute, and atmospheric.)	The gauge pressure relating to atmospheric/absolute pressure
Processes	Responses involving typical chemical engineering process or types of reactors		The new notes that we took today about batches and the way they work
Other Concepts	Responses that relate to material covered in class but that cannot definitively fit into any of the above categories. Most of these responses are only mentioned once.	The formula of volume seems very useful and interesting.	The lecture when we talked about the ideal gas law and what that had to do with what we are talking about

Table 2.3 Week 5 content specific coding categories for reflection prompts

Week 5 (Specific Content)			
Category	Description	Examples	
		Most Surprised	Muddiest Point
DOF	Responses regarding determining degrees of freedom and distinguishing implicit and explicit equations	DOF = Unknowns - species - other independent equations; finally made sense	I learned in lecture that there are implicit and explicit equations that make up the total number of equations accounted for in the degrees of freedom analysis.
Extent of Reaction	Responses regarding the extent of reaction or related concepts	How useful the extent of reaction is in terms of solving the material balances.	zi was kinda confusing, what does it actually represent
Fractional Conversion	Responses relating to fractional conversion. Some included the relationship between fractional conversion and extent of reaction while others involved the relationship between fractional conversion and the amount of remaining moles of a specific species	How intuitive fractional conversions are, most of the work can be derived even without knowledge... key word MOST	how to find the # of moles with a remaining amount
Mass Fraction Eq.	Responses involving the summation of mass fractions being equal to one. This was often double-coded with DOF.	I learned that you can use sum of Xi=1 for each stream instead of only once per subsystem.	
Multiple Streams	Responses regarding processes with multiple streams; these include problems involving recycling and bypass.	bypass	My muddiest point was determining that the composition of a flow split was equal at all parts.

Table 2.3 (continued). Week 5 content specific coding categories for reflection prompts

Week 5 (Specific Content))			
Category	**Description**	**Examples**	
		Most Surprised	**Muddiest Point**
Reactions/MB	Responses focusing on preforming material balances on reactive species	The ability to use chemical equations and relate them to flows.	just in general incorporating the reactions into the balances and stuff. having a rough time picking up concepts and applying them
Rxn Rate	Responses including reaction rate and related principles	That the reaction rate can be generalized to an entire set of species.	
Other Concepts	Responses that relate to material covered in class but that cannot definitively fit into any of the above categories. Most of these responses are only mentioned once.		The equilibrium material.

Results

Word Count Analysis

Figure 2.1 presents a series of boxplots showing the distribution of the word counts of responses by the week of the term and the type of reflection.

The first data shown are from Week 2 since reflections were not administered the first week of the term as it was too close to the start. Week 3 is not shown because data are missing for the Muddiest Point due to a fire drill during that section. Approximately 1,600 student reflections were collected over the term. Students averaged 6 responses with a standard deviation of 3 responses. Within the figure, the dashed boxes are the two sets that were coded in the study. The solid boxes correspond to weeks that there was a midterm exam.

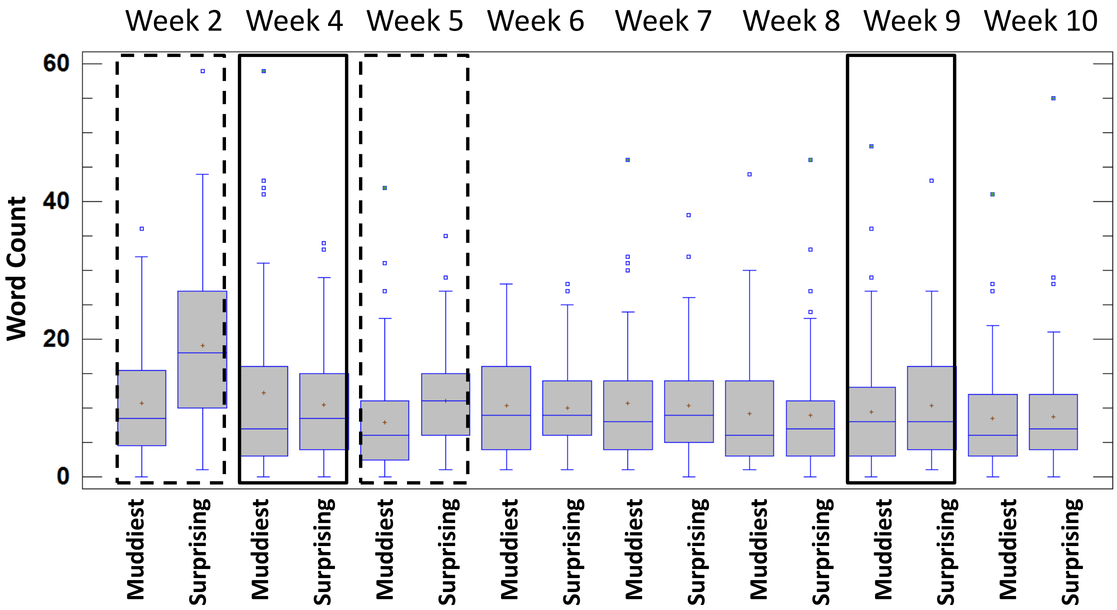


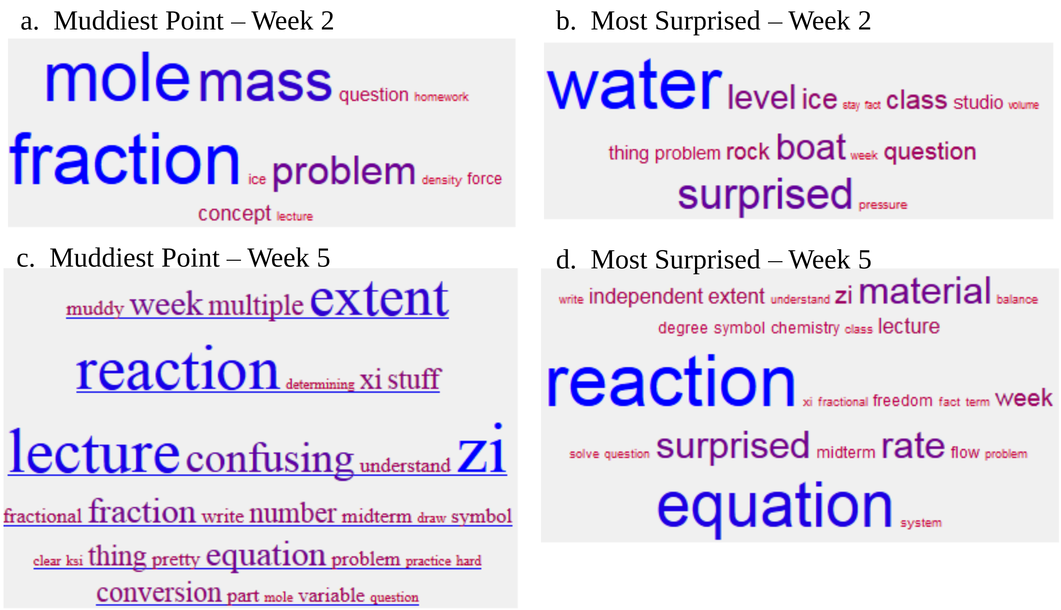
Figure 2.1 Box plots of word count by prompt

One-way ANOVA was used to compare the word count by week. Week 2 (P < 0.001) and Week 5 (P < 0.003) had statistically significantly different counts with the Most Surprised prompt eliciting more words, irrespective of section. It could be interpreted that Most Surprised elicits a higher level of engagement. In Week 2, Section 2 responded to the Most Surprised prompt. In Week 5, Section 1 responded to the Most Surprised prompt.

Midterm exams were administered in Weeks 4 and 9 (shown with solid boxes in Figure 2.1). Correspondingly, the exams became a focus in these weeks with 58 and 78 responses, respectively, that referred to “test”,” exam,” “quiz,” or “midterm,” possibly explaining the similarity in length. Additionally, the difference in word count between prompts reduced as the quarter proceeded. We speculate the students may have started to pay less attention to the instructions as they became familiar with both prompts. Alternatively, data from earlier in the term may simply be anomalous. To further probe differences in responses from the prompts we qualitatively analyzed data from early in the term, described next.

Coding Analysis

Reflection responses from Weeks 2 and 5 were coded using definitions from Tables 2.1-2.3. An overview of the responses in the form of word clouds[21] is shown in Figure 2.2.


a. Muddiest Point – Week 2

b. Most Surprised – Week 2

c. Muddiest Point – Week 5

d. Most Surprised – Week 5

Figure 2.2 Word clouds of Muddiest Point and Most Surprised for weeks 2 and 5

Table 2.4 shows counts for codes that are general to both weeks. In some cases, a response received several codes. Approximately 60% of the responses focused on specific content. There were 35 responses total in the code “Class/HW” early in the term (Week 2) and only 5 by Week 5. While the reference to “Test” were significantly reduced compared to Week 4, there were still 20 responses in Week 5 divided evenly between the two reflection prompts. The Most Surprised prompt elicits responses to pedagogically-oriented reflections not seen in the Muddiest Point responses such as “Misconception” and “Previous Knowledge.” On the other hand, there is a greater reference to “Problem Solving” in the Muddiest Point responses (11 vs. 2).

Table 2.4 General coding counts for reflection prompts

Code Count				
	Week 2		Week 5	
Category	**Most Surprised (Section 2)** n = 134	**Muddiest Point (Section 1)** n = 111	**Most Surprised (Section 1)** n = 100	**Muddiest Point (Section 2)** n = 113
Class/HW	21	14	3	2
Misconception	10	0	0	0
Nomenclature	0	6	10	10

Previous Knowledge	13	0	6	0
Problem Solving	2	6	0	5
Specific Content	86	85	61	74
Test	0	0	10	10
Other	14	4	5	13
Nothing	2	7	5	14

Tables 2.5 and 2.6 show the specific coding counts for Weeks 2 and 5, respectively. In general, the cumulative student reflection show them struggling with similar content with "Buoyancy" receiving the most responses from each prompt in Week 2 and "Extent of Reaction" in Week 5. Clearly these are topics the instructor should provide additional resources for students to learn. The code "Mole/Mass" received a high number of counts for the Muddiest Point in Week 2, but no responses for the Most Surprised. Examination of recitation activity indicates that there was a concept question that asked students to estimate mole fraction of a mixture given the mass fraction. This question was given at the end of Section 1, but there was not time to ask it in Section 2. The "muddiness" of the students with this response may be, in part, due to lack of time to consider the question fully. Instructors should keep contextual factors like this example in mind when interpreting responses.

Table 2.5 Week 2 content specific coding counts for reflection prompts

Code Count		
Category	**Week 2**	
	Most Surprised	**Muddiest Point**
Buoyancy	59	32
Mole/Mass	0	29
Pressure	17	8
Processes	5	6
Application	0	6
Other Concepts	5	4

Table 2.6 Week 5 content specific coding counts for reflection prompts

Code Count		
Category	**Week 5**	
	Most Surprised	**Muddiest Point**
Extent of Reaction	24	40
Reactions/MB	11	9
Fractional Conversion	4	12
Degrees of Freedom	11	3
Multiple Streams	2	6
Reaction Rate	5	0
Mass Fraction. Eq.	4	0
Other Concepts	0	4

Table 2.7 presents coding counts for responses that are clearly positive or negative. In both cases, the majority of responses were coded neutral. However, while the Most Surprised is divided almost equally between positive and negative, the Muddiest Point is more frequently negative. This result has implications to the type of class environment it produces and should be considered by instructors when using reflection prompts.

Table 2.7 Coding counts for positive, negative, and neutral reflection prompts

	Code Count			
Category	**Week 2**		**Week 5**	
	Most Surprised (Section 1)	**Muddiest Point (Section 2)**	**Most Surprised (Section 2)**	**Muddiest Point (Section 1)**
Positive	20	4	25	6
Negative	17	19	25	22
Neutral	98	88	49	85

Specific Responses

We consider next responses for the most common content-specific category from Week 2, buoyancy. For context, in the recitation that the reflections were collected, a conceptual question was asked in which the students were asked to predict how the level of liquid water compares for a system initially composed of ice water to the level of the same system after the water melts. The normatively correct response recognizes that the ice is initially partially submerged at the water interface and applying the principle of buoyancy reasons that there is no change in water level after it melts.

The following shows Most Surprised and Muddiest Point responses that appear to refer to that question. They are selected for a specific case, but are representative of the general type of differences we observed for each prompt. For the Most Surprised prompt, students responded:

- "…The ice being the same volume when melted"
- "The ice cubes melting causing the water level to stay the same"
- "The fire! Not actually, honestly, it was this question from today about the level of water and ice---initially puzzling, but then it makes sense in principle!"

From the buoyancy-coded responses to Muddiest Point, students responded:

- “The problems that involve buoyant force like the one with melted ice.”
- “The rock on the boat problem and the ice cube problem in recitation. Having difficulty figuring out the change in water level when objects are added/removed/submerged in a liquid”
- “The image in the ice melting question was misleading because the ice appeared to be fully submerged.”

From an instructor’s perspective, it is harder to recognize what action to take with the Most Surprised responses [the instructor might interpret the first response as conceptually incorrect (or perhaps just poorly worded)]. In the third response, the student is clearly celebrating her/his conceptual processing and perceived gains in understanding. The first two responses from Muddiest Point provide more direction for the instructor. The third response might be interpreted as more focused on the representation in the concept question than on an understanding relative to the concepts that week.

Instructor Classroom Response to Reflection Data

The instructor provided a summary of how he used these reflection data in class. He expected that students would reflect primarily on the in-class activities. The in-class structure includes two lectures, one studio, and one recitation (where Concept Warehouse conceptual questions are answered). However, responses often referred to homework assignments and exams as well. As shown in Table 2.4, responses in a given week covered a broad range of topics. Sometimes a single challenging concept could be inferred from the large number of responses such as a question from the previous hour of recitation and sometimes responses referenced a concept introduced several days prior.

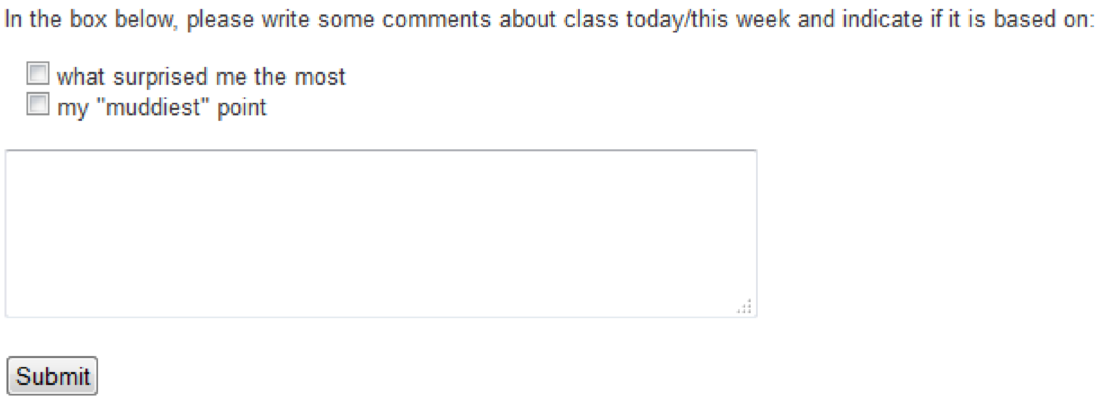
Most weeks, the instructor reviewed the responses in between the recitation where they were collected and the next lecture and identified common themes. In a class with nearly 300 students, reading student responses was time consuming and the time the instructor had available to review responses varied. However, the instructor thought that the review was always valuable and always identified one or more topics were to explicitly address in the next class period. This practice was felt to be critical to “closing the loop” with respect to the student feedback. By explicitly acknowledging student responses and taking corrective action, the instructor believed that students would feel they had a voice and motivated to speak candidly about aspects of the material with which they were struggling. Corrective action was generally quite simple, ranging from a few clarifying statements and suggestions for additional resources to review of previous material and additional example problems.

In reflecting on this implementation there are several changes that the instructor would recommend. First, the instructor thought it better to choose either the Muddiest Point or Most Surprised reflection and use that the entire term. He believed students started to view these as the same over time when the questions were alternated. Second, efforts should be made to ensure sufficient time (3-5 minutes) for thoughtful completion of this activity. In several instances, the activity was too rushed at the end of the class period. It became clear in reviewing the responses that if an instructor would like useful information about confusing items in a specific element of the course, the questions should be deployed with that specific intent (e.g., at the end of each class period). Ideally, there would be feedback on each aspect of the course. One could imagine a continuous dialog where these questions are asked at the end of each activity with the expressed intent of responding to the feedback in the subsequent class period. Finally, the instructor would make better use of the built-in tools in the *Concept Warehouse* (e.g., clickable word clouds) to quickly drill down into the most common responses. With a better handle on the common themes, the instructor can more effectively address the feedback during class and add value to the exercise.

Discussion

The Muddiest Point and Most Surprised reflection activities serve two purposes. First, they provide information and communication to the instructor about the attitudes, understanding, and learning approaches of the students. This aspect allows an instructor to directly and immediately address the specific difficulties and concerns that arise. It also provides the instructor information such as how to better align upcoming content with prior knowledge. Second, the activities encourage students to reflect and be metacognitive about their own learning. Here students consider and evaluate their own learning relative to the course objectives, processes, and structures. In this aspect, it is important to consider the different ways that the language of each prompt positions the student and to consider the affective responses that may be elicited.

In balancing these purposes, it appears that each prompt has benefits. The Muddiest Point is more familiar to students and directly asks students to identify concepts that are confusing. However, Most Surprised doesn't necessarily have a negative connotation and can provide opportunity for students to reflect on and express their successes. So what should an instructor do? Of course, the answer depends on context. One response could be to do both, taking advantage of the affordances of each. Based on the findings and analysis, we suggest an alternative, shown in Figure 2.3. This alternative provides students the opportunity to select Muddiest Point, Most Surprised, or both. Such a strategy provides authorship to the student. It also allows an alternative prompt to those students who wish to express a positive outlook. If the class size is large enough, as the case studied here, the instructor should have sufficient Muddiest Point responses to directly identify unclear and confusing content, but also the broader pedagogical and affective responses seen in Most Surprised. We plan to implement this approach in the same class next year and study the responses.



In the box below, please write some comments about class today/this week and indicate if it is based on:

☐ what surprised me the most
☐ my "muddiest" point

Figure 2.3 Second generation reflection activity envisioned based on the results and analysis in this study

Acknowledgements

The authors are grateful for support provided by the National Science Foundation grant DUE 1347817. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

1. Kaplan, M., Silver, N., Lavaque-Manty, D., & Meizlish, D. (2013). *Using reflection and metacognition to improve student learning: Across the disciplines, across the academy.* Sterling, VA: Stylus Publishing.
2. Bransford, J., A. Brown, and R. Cocking. *How People Learn: Brain, Mind, Experience and School*. Washington, D.C.: Commission on Behavioral and Social Science and Education, National Research Council. 2000.
3. Consortium to Promote Reflection in Engineering Education, http://depts.washington.edu/celtweb/cpree/, accessed 02.01.2015
4. Mosteller, F. (1989). The "muddiest point" in lecture as a feedback device. *On Teaching and Learning: The Journal of the Harvard-Danforth Center*, 10-21. http://isites.harvard.edu/fs/docs/icb.topic771890.files/OTL3-Mosteller-Muddiest.pdf
5. Angelo, T. A., & Cross, P. K. (1993). Classroom assessment technique examples. In Classroom Assessment Techniques: A Handbook for College Teachers (2nd ed.) Retrieved from http://www.ncicdp.org/documents/Assessment%20Strategies.pdf
6. Hall, S. R., Wait, I., Brodeu, D. R., Soderholm, D. H., & Nasr, R. (2002). Adoption of active learning in a lecture-based engineering class. *Frontiers in Education.* doi: 10.1109/FIE.2002.1157921
7. Tanner, K. D. (2012). Promoting student metacognition. *CBE—Life Sciences Education 11*, 113–120. doi: 10.1187/cbe.12-03-0033
8. Krause, S. J., Baker, D. R., Carberry, A. R., Koretsky, M., Brooks, B. J., Gilbuena, D., Waters, C. & Ankeny, C. J. (2013). Muddiest Point Formative Feedback in Core Materials Classes with YouTube, Blackboard, Class Warm-ups and Word Clouds. In *Proceedings of the 2013 American Society for Engineering Education Annual Conference & Exposition*.
9. Krause, S. J., Baker, D. R., Carberry, A. R., Alford, T. L., Ankeny, C. J., Koretsky, M., Brooks, B. J., Waters, C., Gibbons, B. J., Maass, S., Chan, C. K. (2014) Characterizing and Addressing Student Learning Issues and Misconceptions (SLIMs) in Materials Science with Muddiest Point Reflections and Fast Formative Feedback. In *Proceedings of the 2014 American Society for Engineering Education Annual Conference & Exposition*.
10. King, D. B. (2011). Using clickers to identify the muddiest points in large chemistry classes. *Journal of Chemical Education* 88(11), pp 1485-1488. doi: 10.1021/ed1004799
11. Brookfield, S. D. (1995). *Becoming a Critically Reflective Teacher*. San Francisco: Jossey-Bass.

12. Brookfield, S. (1998). Critically reflective practice. *Journal of Continuing Education in the Health Professions*, *18*(4), 197-205.
13. Glowacki-Dudka, M., & Barnett, N. (2007). Connecting Critical Reflection and Group Development in Online Adult Education Classrooms. *International Journal of Teaching and Learning in Higher Education*, *19*(1), 43-52.
14. Adams, K. L. (2001). The Critical Incident Questionnaire: A critical reflective teaching tool. Exchanges:The Online Journal of Teaching and Learning in the CSU.
15. Hessler, H. B., & Taggart, A. R. (2011). What's stalling learning? Using a formative assessment tool to address critical incidents in class. *International Journal for the Scholarship of Teaching and Learning*, *5*(1), 9.
16. Hedberg, P. R. (2009). Learning through reflective classroom practice. Journal of Management Education, 33(1), 10-36.
17. Boyle, J. T., & Nicol, D. J. (2003). Using classroom communication systems to support interaction and discussion in large class settings. Research in Learning Technology, 11(3).
18. Brookfield, S. D., & Preskill, S. (1999) *Discussion as a way of teaching* (Vol. 85). San Francisco: Jossey-Bass.

3. SURPRISES IN THE MUDDY WATERS OF HIGH-ENROLLMENT CLASSES

Jessie Keeler & Milo Koretsky

Introduction

Classroom assessment techniques that ask students to reflect on material covered in class are believed to help improve learning by allowing the student to more actively participate in the learning process through evaluating their understanding of course content.[1] Promoting students to be more reflective about their learning experiences also allows them to develop robust learning strategies and metacognitive skills that are characteristic of expertise.[2] In addition, students' written reflections can provide instructors with formative information about students conceptions of course content. In this paper, we use the definition of Turns and Atman that reflection is the act of "exploring the meaning of experiences and the consequences of the meanings for future action."[3]

In a common type of reflection activity, often termed the minute paper[4-6] or the exit slip,[7] students write a short reflection on the content they have just learned in class. In a previous study,[8] we compared two end of class reflection questions which we labeled Muddiest Point and Most Surprised. The Muddiest Point required the students to respond to the prompt, "What was the muddiest point in class this week?" whereas the Most Surprised activity asked students to answer the question "What surprised you most about class this week?" In that study, we used a quasi-experimental design to examine students' in-class responses to these two exit questions where students were asked to reflect on the class over the last week. Each week, one section of students was given the Muddiest Point prompt while the alternate section was given the Most Surprised prompt.

Based on the findings and analysis of the previous study, we concluded that there are benefits specific to each prompt. The Muddiest Point prompt is more familiar to students and it easily allows them to directly express what course content is unclear. The Most Surprised prompt allows a wider range of responses. In this article, we report on our continued examination of student reflection responses through implementation of a hybrid prompt that allows students the opportunity to select Muddiest Point, Most Surprised, or both. This strategy provides greater authorship to the student. It also allows an alternative prompt to those students who wish to express a positive outlook. If the class size is large enough, as are the cases studied here, the instructor should have sufficient Muddiest Point responses to directly identify unclear and confusing content, but also the identification and expression of student success seen in Most Surprised.

In this research study, we analyze the implementation of the hybrid prompt in two different large courses (around 200 students each) taught by different instructors. We ask the following research questions:

RQ1. How frequently do students choose each type of prompt (Muddiest Point, Most Surprised, or both)? Do their choices tend to change during the term?

RQ2. Does the type of content in the student reflection responses differ based on the type of prompt that they choose? Does the type of content tend to change during the term?

RQ3. Is there a difference in the affective quality of the responses for the Muddiest Point and Most Surprised prompts?

Background

Instructors have used several short end-of-class reflection activities to both promote student learning and gather formative information about students' conceptions of course content. Prompts include requesting students to describe the most important,[9-11] interesting,[12] confusing or muddiest,[8,10-15] and surprising[8,16] aspects from a lecture, recitation, or week of the course. In this paper we focus on the latter two.

Some of the early uses of the Muddiest Point activity focused on providing formative information for the instructor. For example, Mosteller[5] reported on the use of the Muddiest Point activity in a statistics class and argued that the practice "…promotes concrete and sometimes non-trivial responses to the question of what 'you want to know more about'"(p. 19). A few years later Angelo and Ross[6] described the muddiest point as having one of the best returns on investment;

although very little time is required to perform a Muddiest Point implementation, it can yield useful information about student understanding.

More recent studies emphasize the use of the Muddiest Point activity to promote student learning, particularly metacognition. For example, Tanner[8] found that assigning a Muddiest Point activity helps students realize that being confused is a natural aspect of learning. This activity also was found to allow the students to safely share what they find unclear without having to reveal their confusions to the entire class and thus risk judgment. Tanner argued that while students might be unaccustomed to being open with their professors about their confusions, the process of acknowledging, embracing, and resolving their misconceptions is a key to evolving as learners.

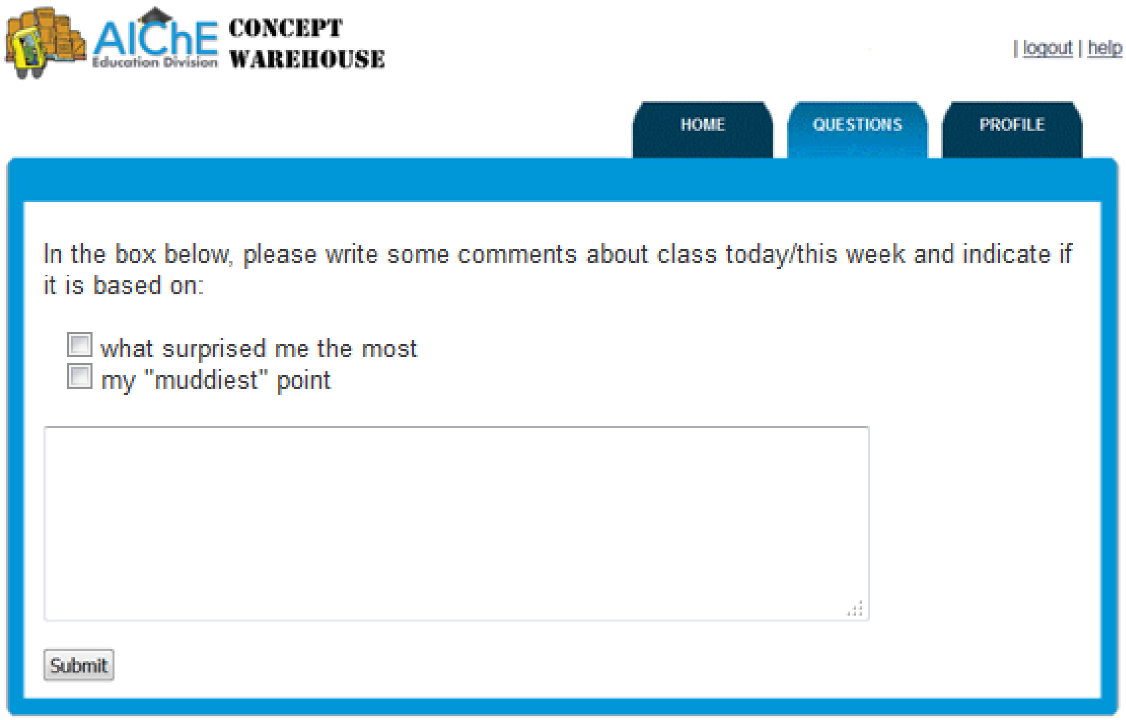
Hall and colleagues[7] report on a study in the context of increasing active learning in an engineering class. That study incorporated activities such as concept tests, group discussions, and Muddiest Point evaluations across several different engineering disciplines. They found that a majority of professors and students found Muddiest Point activities to be effective. Students felt that the professors' clarifications about the Muddiest Points directly improved their learning, showed their professors cared, and enhanced their overall relationship with their professors. The exercise was also popular amongst professors; several planned to include the exercise in their future courses. In Introductory Materials courses, Krause and colleagues[9,10] found the use of Muddiest Point activities informed instructors' use of formative process feedback and improved student attitudes, achievement and retention of course content.

Most Surprised activities are rarely used in science, but have been used by instructors in other fields in a way similar to Muddiest Point. In the most common form reported in the literature, Most Surprised is posed as one question in Brookfield's[12] five-question critical incident questionnaire (CIQ). The CIQ is intended to be assigned to students at the end of a course period or week of a course, in a similar manner to the Muddiest Point activity. In the CIQ, Most Surprised comes as the fifth and final item and is proceeded by questions regarding when students felt most engaged and least engaged, and what actions taken in class were most confusing and most affirming.

The use of the CIQ has been reported in many different educational settings, including health education,[17] adult online education,[18] public communication,[19] writing,[20] management education,[21] and engineering mechanics.[22] Generally the CIQ has been described as a useful instructional tool that promotes learner reflection. However, studies that explicitly assess the tool's effectiveness in various settings are lacking.[23]

Methods

In this article, we report on student responses to a modified reflection prompt that provides more authorship to the student by allowing her/him to choose either reflection or both. This activity is available through the AIChE Concept Warehouse[24] and shown in Figure 1. We conducted the study at a large public university in the Northwest United States. Participants were enrolled in either a sophomore-level course in Process Analysis (2nd Year Course) or a junior-level course in Chemical Thermodynamics (3rd Year Course). Each course was taught by a different instructor. Data are only reported for students who agreed to participate and signed an informed consent form approved by the Institutional Review Board.



In the box below, please write some comments about class today/this week and indicate if it is based on:

☐ what surprised me the most
☐ my "muddiest" point

Figure 3.1 Hybrid prompt reflection activity used in this study

The Concept Warehouse allows students to respond in class using their smart phones or other devices. However, both instructors chose to have the students respond to the reflection prompts outside of class and allowed around one week for students to submit their reflections. Both instructors read through the student reflections weekly and communicated to the students about them. The instructor in the 2nd Year Course responded to the reflections in writing and posted the responses on the learning management system for the course. The instructor for the 3rd Year course selected representative responses and displayed them to the class and led an interactive discussion.

For each course, we analyzed three sets of student responses using an emergent coding scheme. We wanted longitudinal data to evaluate if student responses changed during the term. We also wanted to avoid coding weeks that aligned with test dates. Based on these criteria, we coded responses from Week 1, Week 3, and Week 7 for the 2nd Year Course and Week 1, Week 4, and Week 7 for the 3rd Year Course. Table 1 shows the number of responses coded for each week for each course. To member-check the analysis and to better understand how the instructors made use of the reflection activity, we regularly conducted non-scripted interviews with the instructors during the term. The interviews generally inquired about their experience with the activity and their approach to addressing student responses. We also asked about their attitudes towards the hybrid prompt relative to the more common Muddiest Point by itself.

Table 3.1 Number of students responding each week

Week	**Number of coded responses**	
	2nd Year Course	**3rd Year Course**
Week 1	192	183
Week 3, 4	178	185
Week 7	152	176

To answer Research Question 1, we counted how many times students chose each prompt (Muddiest Point, Most Surprised, or both). To answer Research Questions 2 and 3, we analyzed responses using open coding, a process used to infer categories of meaning. The coding process involved reviewing the responses and sorting them into categories. Initially each set of data was scanned for reoccurring responses that could be used to generate the coding categories. After the first set of categories were created, there were a few responses left without a category. If these responses could be related to one another, a new category was created for them. When all of the remaining responses were unconnected both to each other and to existing categories, they were coded with the catch-all category "Other." Once coding categories were established, fifty responses for each prompt (100 total) were independently coded by two researchers. An interrater reliability using the Cohen's Kappa statistic of 0.94 for Muddiest Point and 0.92 for Most Surprised was achieved. These values show acceptable reliability for the coding process.

This assessment allowed us to categorize the content and affect of the student responses. We then quantified these data by recording the number of codes in sets of categories. Since the resulting data are categorical, we used Chi-square tests for independence. For cases with two variables, analysis used contingency tables while for more than two variables, log linear (Poisson) regression was used. A threshold of $\alpha = 0.05$ is used to determine significance.

Overall, 17 and 19 coding categories were generated for the 2nd and 3rd year courses, respectively. From those codes generated we focus on two aggregate categories: “Structure” and “Content.” Codes contained in the Structure category involve elements of how the course was organized and delivered; they do not contain references to specific course content. These codes include “teaching style,” “course structure,” “studio,” “book,” and “homework.” The Content category consists of codes related to course-specific conceptual knowledge and procedural fluency through codes including “course content,” “example,” and “specific problem.” Table 2 and Table 3 contain a list of the generated Structure and Category codes, respectively, as well as a code description, and both a Muddiest and Most Surprised example response. For simplicity and ease of comparison between cases, the examples are all taken from the 3rd Year Course. Table 4 shows the coding scheme for affective responses. When a response was either clearly “Positive” or clearly “Negative” it was coded as such; otherwise, it was considered neutral.

Table 3.2 Structure coding categories for reflection prompts

Structure Codes			
Category	**Description**	**Examples (from the 3rd Year Course)**	
		Most Surprised	**Muddiest**
Teaching Style	Responses involving the instructor, their teaching approach and/or how they run the class	What surprised me the most were the simple and clear explanations of equilibrium and other thermodynamic concepts. For example, I felt [instructor] clearly explained the relationship between an extensive variable and intensive (molar or specific) variable with the equation K=k*n…	I feel that following Friday's lecture was too confusing. I did not understand the way that instructor writes the equations and he he ended up with fact. I needed to go back to the book and check out the details of that equations and what the symbols are representing for.
Class Structure	Responses relating specifically to the organization and framework of the course	I was surprised at how the class is structured. I like participating in class on Wednesdays via the concept warehouse site through answering interactive questions. I like that this term there will be more of a focus on the explanation for these concept warehouse questions because often times I feel as though I know an answer but I am not positive as to why it is correct.	My muddiest point this week was the connection between studio and class. We were working with liquid vapor equillibrium in studio and gibbs free energy, while in lecture it has been more of a review of 311. I guess im just having a hard time getting into the swing of things.
Studio	Responses focusing mainly on the studio session of class in which students break into small groups and complete worksheets	The studio simulation, was tricky to figure out but was a pretty cool simulation once we got the idea of how to run it.	My muddiest point was during studio. We were only given one worksheet per team. We completed both the studio and the homework part A section, but I'm unsure whether or not I need to fill out another one individually or not…
Homework	Responses relating to the overall role homework plays in the course as opposed to specific concepts within homework assignments	That we turn in 2 HWs per week	The muddiest point is that we have too much homework. So it is hard to manage our time and give some time for the other homework.
Book	Responses focusing on the role the textbook plays in the course	There is a gap between the text and the class	Why the people that purchased the physical copy of your book get an unfair advantage over the people that purchased the online edition on exams. The people with the physical books can write as much as they'd like in the margins of their books, but the people that had to have "pre-checked" pdf copies did not have that advantage. Open book = Open note.

Table 3.3 Content coding categories for reflection prompts

Structure Codes			
Category	**Description**	**Examples (from the 3rd Year Course)**	
		Most Surprised	**Muddiest**
Example	Responses involving sample problems or demonstrations done in class	I've known since physics 211 that volume is not conserved but it was really surprising to see it change so much when mixing two different liquids. 2mL out of 80mL is over 2%. I wouldn't have guess it would change in a way that you could measure with the naked eye.	The cartoon representations of a and b were very hard to follow.
Specific Problem	Responses focusing on particular problems students worked through in the course	What surprised me was when two equal sized compartments of O2 was separated by a partition, and when the partition was removed, there was no change in the entropy of the system.	How did the equation from studio with the ln(P2/P1) and ln(T2/T1) turn into one with only mole fractions? Also, why are intermolecular interactions more important than configurations when finding change in entropy?
Course Content	Responses relating to specific content covered in class	What surprised me most about this week was how an increase of pressure meant a decrease in entropy because I have never thought about the relation of pressure to entropy change in my previous courses.	I am still having troubles understanding how Gibbs energy changes when we have a real gas versus having an ideal gas. Also I do not completely understand partial molar gibbs energy.

Table 3.4 Affective coding categories for reflection prompts

Affective Codes			
Category	**Description**	**Examples (from the 3rd Year Course)**	
		Most Surprised	**Muddiest**
Positive	Responses with clearly positive connotations.	I really liked the format of the group participation Concept Warehouse on Wednesday. It was a much more helpful method of understanding the base information. Much better than any other class I've been to that has used Concept Warehouse. I think I'm finally nailing down the concept of entropy.	After the class this week I have a better understanding of the relationship between Gibbs free energy and equilibrium, and how to judge the direction of phase transfer based on certain equations.
Negative	Responses with clearly negative connotations.	How terribly structured the studio is. The TA said they can't help us. What's the point of even having a studio if we get no feedback or at best late feedback a week later. That isn't helpful at all.	I had a lot of trouble in the studios. I know the point is to kind of leave us with minimal guiding in the beginning, but my studio TA was the opposite of helpful. I had the correct answer on my sheet, the TA told me something incorrect so I changed the answer, then he came back and indicated I had the wrong answer and had indeed been correct initially, and it was a huge waste of time.
Neutral	Responses not coded as positive or negative.	The part about lecture that surprised me the most was that at equilibrium the change in moles of a substance at state alpha is equal to negative of the change in moles at state beta.	I was confused on how volume may relate to Gibb's energy. Also, I wasn't sure when to use the definition itself or the differential form of Gibb's energy when analyzing situations.

Results

The percentage of students who chose each prompt is shown by class and week in Table 5. In every case, more students chose to respond to the Muddiest Point prompt (50% - 75%) than the Most Surprised prompt (20% - 42%), and relatively few students chose to use both prompts (2% - 8%). As the term progressed, the percentage of students that replied to the Muddiest Point prompt relative to the Most Surprised significantly increased for both the 2nd Year Course ($\chi^2 = 23.3$, df = 6, p = 0.0007) and the 3rd Year Course ($\chi^2 = 24.5$, df = 6, p = 0.0004).

Table 3.5 Percentage of students responding to each reflection prompt

Week	Muddiest Point		Most Surprised		Both	
	2nd Year Course	3rd Year Course	2nd Year Course	3rd Year Course	2nd Year Course	3rd Year Course
Week 1	59%	50%	36%	42%	5%	8%
Week 3, 4	74%	61%	24%	32%	2%	8%
Week 7	75%	67%	20%	27%	5%	6%

We next present results to open coding analysis. As explained above, most of the generated coding categories were grouped into either “Structure” or “Content” aggregate categories (See Tables 2 and 3, respectively). The coding category “other” was not included in either aggregate category. The code counts were distributed broadly across the various Structure categories in Table 2. In contrast, the majority of code counts for the Content category in Table 3 were identified with the “Course Content” code and were specific to the topic of that given week. For example, in Week 4 in the 3rd Year course, Course Content coded reflections referred to concepts of “entropy,” “hypothetical paths,” “partial molar properties,” and “Gibbs Energy.”

Table 6 contains data of the Content and Structure codes from the 2nd Year course and Table 7 contains data from the 3rd Year course. Log linear regression analysis shows that there is a significant change in the dependent variable (Structure vs. Content) with both week ($\chi^2 = 14.8$, df = 2, p = 0.0006) and with prompt ($\chi^2 = 15.9$, df = 3, p = 0.0012), but there is not a significant difference with course ($\chi^2 = 0.14$, df = 1, p = 0.712). As the term progressed in both courses, significantly more student responses were directed towards the content (concepts and procedures) of the course rather than its structure. In addition, students were more likely to reflect on structure through the Most Surprised prompt. Even though the courses were at different levels and taught by different instructors who communicated the results of the activity differently, there was not a significant difference in these trends between courses.

Table 3.6 Percentage of aggregate code by prompt for 2nd year course

2nd Year Course	Content		Structure		Both	
	Muddiest Point	Most Surprised	Muddiest Point	Most Surprised	Muddiest Point	Most Surprised
Week 1	43%	9%	15%	26%	4%	3%
Week 3	49%	9%	24%	8%	9%	2%
Week 7	74%	17%	2%	2%	5%	1%

Table 3.7 Percentage of aggregate code by prompt for 3rd year course

3rd Year Course	Content		Structure		Both	
	Muddiest Point	Most Surprised	Muddiest Point	Most Surprised	Muddiest Point	Most Surprised
Week 1	37%	25%	14%	14%	5%	5%
Week 4	59%	21%	5%	7%	5%	3%
Week 7	67%	13%	6%	11%	2%	1%

Table 8 shows the percentage of clearly positive or clearly negative affective responses. Two responses were coded as both positive and negative and although we included them in the statistical analysis, they are omitted from Table 8. Since the majority of responses were neutral, we report the percentages relative to the total positive and negative codes for each course; however, the statistical analysis includes neutral responses as well. The Most Surprised prompt, which students can use to express successes, elicits significantly more positive responses while the Muddiest Point prompt elicits more negative responses ($\chi^2 = 22.8$, df = 3, $p < 0.0001$). The number of positive responses also decreases with week ($\chi^2 = 26.4$, df = 2, $p < 0.0001$). Again there is not a significant difference between the 2nd Year Course and the 3rd Year Course ($\chi^2 = 2.3$, df = 1, $p = 0.130$).

Table 3.8 Percentage of affective coding

Prompt	Positive		Negative	
	2nd Year Course	3rd Year Course	2nd Year Course	3rd Year Course
Muddiest	7%	7%	39%	31%
Most Surprised	38%	42%	12%	10%
Both	1%	3%	2%	6%

Discussion

Muddiest Point and Most Surprised reflection activities have several purposes. First, they communicate information to the instructor in regards to the attitudes, understanding, and learning approaches of the students. The specific difficulties and concerns that emerge can then be immediately addressed. With these reflection activities, the instructor also gains insight into aligning upcoming content with prior knowledge for better levels of comprehension. Second, these activities foster metacognitive and reflective awareness in students. They must contemplate and gage their own learning relative to the course objectives, processes, and structures.

Taking Reflection beyond Misconceptions

A topical focus such as responses identified in our "Content" category is typical of the uses of Muddiest Point reflections reported in the literature.[5, 9-11] In this way, students communicate to instructors and to themselves topic-specific concepts and procedures that are difficult for them. We see this type of information provided to instructors in the responses in this study, and the proportion of this type of response tends to increase as the course progresses. However, the use of the reflection activity shown in Figure 1 also provided the instructor feedback on course structure, information not typically identified with Muddiest Points alone. We found that many students used the reflection questions to submit suggestions for ways to improve the effectiveness of course management and delivery. In this way, the reflection activity functions in the manner of a regular (weekly) formative student evaluation of teaching. An instructor who reviews responses can adjust the structure of the course or their teaching practice without the need of a formal mid-term evaluation and without waiting until the end of course evaluation. The decrease in the proportion of responses categorized with the Structure code may indicate that changes were being made. We next provide two specific examples.

The following response was available to the Instructor after Week 1 in the 3rd year course:

- **Muddiest Point:** In class, the notes [Instructor X] writes on the whiteboard are not large enough to be seen in the back of the room. I think a better way to present notes in class would be to write on paper and show it on the DOC CAM.

This Muddiest Point response was one of several that the instructor received about the difficulty of reading what he wrote on the whiteboard during class. Historically, this instructor had used a white board or chalk board during class. However, increasing enrollments had shifted the course to larger and larger classrooms. What worked well with a smaller class (writing on the board) had apparently become troublesome for those students who were further away. These responses provided helpful and immediate feedback that led to the instructor adjust how he presented material. While this adjustment was simple to make, the reflection activity provided communication that may not have occurred otherwise.

In the same week, the instructor read the following response

- **Most Surprised:** I really liked the format of the group participation Concept Warehouse on Wednesday. It was a much more helpful method of understanding the base information. Much better than any other class I've been to that has used Concept Warehouse. I think I'm finally nailing down the concept of entropy.

The instructor had deliberately framed the Concept Warehouse activities by discussing with the class characteristics of scientific reasoning and collaborative meaning making. The Most Surprised reflection provided information to the instructor that his intentional emphasis on collaborative reasoning during active learning was taken up (at least by this student).

Many responses suggested conceptual difficulty or structural deficiencies. In their interviews, both instructors commented that having a set of positive or even personal responses (such as when a 3rd Year Course student submitted: “most surprised by how much you like coffee”) interspersed uplifted them. Reading about aspects of class that students enjoyed or the successes students celebrated was encouraging to the instructors.

Research Questions

We next address the research questions for this study:

For Research Question 1, we found that more students selected the Muddiest Point prompt, but still a significant proportion choose the Most Surprised prompt, and a small amount selected both. As the term progresses, the proportion of Muddiest Point responses increases and Most Surprised decreases.

For Research Question 2, we found that both prompts elicited responses about both the content and the structure of the course; however, the proportion of responses was different. Students who chose the Muddiest Point prompt were more likely to contain content specific responses while the Most Surprised responses were split evenly between content and course structure. As the term progresses students focus more on course content and problematic concepts instead of commenting on the overall class organization or the teaching practices of the instructor. This shift can be explained by adjustments of class participants; as students adjust to the way the class is taught and instructors adjust course delivery based on student feedback, the students become more likely to reflect on specific course content as opposed to course structure. This explanation is consistent with the results from Research Question 1 which show that the Muddiest Point prompt becomes more prominent throughout the term. Alternatively, it is plausible that a greater focus on content results since course content becomes more sophisticated as the term progresses.

For Research Question 3, we found that although the majority of the responses from the two courses were coded as neutral, there was a correlation between the two reflection prompts and the responses that were coded as positive or negative. The Most Surprised prompt yielded more positive responses whereas the Muddiest prompt was more frequently negative. The Most Surprised prompt allows students an opportunity to reflect on their successes throughout the course; these positive responses are also uplifting for instructors as they review student responses.

Limitations

This study has several limitations. The reflection activity was only implemented at one university and with one group of students. Although two courses were involved, the data from the 3rd Year Course was gathered the year following the 2nd Year Course; therefore, it was collected from the same cohort and there was significant overlap in the group of students sampled. To determine the generalizability of the results presented in this article, similar implementations of this hybrid reflection activity and analyses of student responses should be done at other institutions, in other disciplines, and with other groups of students. In order to encourage student engagement in the reflection activities, both instructors offered participation points to students who submitted responses. While this practice encouraged student responses, students who were only concerned with their grades may have submitted lower quality reflection responses. Finally, there is inevitable inference in determining positive or negative connotations from student writing. To address this concern, a conservative coding protocol was developed where the researchers only coded responses as positive or negative when it was abundantly clear the student projected those emotions; however, this coding process led to most of the responses being coded as neutral.

The instructors also experienced some limitations while implementing the reflection activity. The instructor of the 2nd Year Course commented that responses that would be coded as “Structure” could be contradictory, and he was unsure how to respond. He also expressed disappointment by the lack of depth of some student responses. However, the 3rd Year Course instructor did not report any concerns about the ambiguity of structure or the quality of student responses. While both instructors read the reflections and responded, they had very different practices in communicating with the class. The 2nd Year Course instructor posted the responses on a learning management system while the 3rd Year course led an interactive discussion in class. We also saw a significant decrease in the student participation in the reflection activity in the 2nd Year Course but not in the 3rd Year Course (Table 1). We speculate all these factors may be related and suggest research connecting classroom practice around communication in this type of reflection activity is needed.

Implications

In this study, we analyzed student responses to a hybrid reflection activity where students could select amongst a Muddiest Point, a Most Surprised, or both. It is well established that Muddiest Point reflections activities provide instructors with feedback on the conceptual and procedural challenges students are facing.[6,10] While our results confirm these findings, we also noticed students were able to report structural barriers. There is increasing interest in formal formative assessments of teaching during a class (as opposed to the end). The information in our hybrid reflection allows instructors a sense of what does or does not work for students early in a course, and regularly throughout the course, as opposed to relying on typical end of the term teaching assessments. Our study suggests that this type of information can be obtained by appropriate reflection activities and does not need to be an additional midterm evaluation survey. Receiving weekly feedback allows instructors to adjust classes as they're occurring instead of waiting until the next iteration of the course.

We also found that the hybrid activity more readily allowed positive experiences to be expressed and shared mostly through the use of the Most Surprised prompt. We believe opportunities for such positive communication make for a healthier learning ecosystem for both students and the instructor.

Acknowledgements

The authors are grateful for support provided by the National Science Foundation grant DUE 1347817. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

1. Kaplan, M.; Silver, N.; Lavaque-Manty, D.; Meizlish, D. (2013). *Using Reflection and Metacognition to Improve Student Learning: Across the Disciplines, Across the Academy;* Stylus: Sterling, VA, 2013; pp 78-103.
2. Bransford, J.; Brown, A.; Cocking, R. *How People Learn: Brain, Mind, Experience and School*; National Academy Press: Washington D.C., 2000; pp 12-20
3. Consortium to Promote Reflection in Engineering Education, http://depts.washington.edu/celtweb/cpree/ (Accessed February 1, 2015).
4. Keeler, J.; Brooks, B. J.; Friedrichsen, D. M.; Nason, J. A.; Koretsky, M. What's Muddy vs. What's Surprising? Comparing Student Reflections About Class. In *2015 ASEE Annual Conference*, Proceedings of the 2015 American Society for Engineering Education Annual Conference & Exposition, Seattle, WA, June 14-17, 2015.
5. Mosteller, F. The "Muddiest Point" in Lecture as a Feedback Device. *On Teaching and Learning: The Journal of the Harvard-Danforth Center*, **1989**, 10-21.
6. Angelo, T. A.; Cross, P. K. *Classroom Assessment Techniques: A Handbook for College Teachers;* Jossey-Bass: San Francisco, CA, 1993; pp 154-158.

7. Hall, S. R.; Waitz, I.; Brodeu, D. R.; Soderholm, D. H.; Nasr, R. Adoption of Active Learning in a Lecture-Based Engineering Class. In *Frontiers in Education*, Proceedings of the ASEE/IEEE Frontiers in Education Conference, Boston, MA, November 6-9, 2002.
8. Tanner, K. D. Promoting Student Metacognition. *CBE Life Sci. Educ.* **2012,** *11*, 113–120.
9. Krause, S. J., Baker, D. R., Carberry, A. R., Koretsky, M., Brooks, B. J., Gilbuena, D., Waters, C. & Ankeny, C. J. Muddiest Point Formative Feedback in Core Materials Classes with YouTube, Blackboard, Class Warm-ups and Word Clouds. In *2013 ASEE Annual Conference,* Proceedings of the 2013 American Society for Engineering Education Annual Conference & Exposition, Atlanta, GA, June 23-26, 2013.
10. Krause, S. J.; Baker, D. R.; Carberry, A. R.; Alford, T. L.; Ankeny, C. J.; Koretsky, M.; Brooks, B. J.; Waters, C.; Gibbons, B. J.; Maass, S.; Chan, C. K. (2014) Characterizing and Addressing Student Learning Issues and Misconceptions (SLIMs) in Materials Science with Muddiest Point Reflections and Fast Formative Feedback. In *2014 ASEE Annual Conference*, Proceedings of the 2014 American Society for Engineering Education Annual Conference & Exposition, Indianapolis, IN, June 15-18, 2014.
11. King, D. B. Using Clickers to Identify the Muddiest Points in Large Chemistry Classes. *J. Chem. Educ.* **2011,** *88*(11), 1485-1488.
12. Brookfield, S. D. *Becoming a Critically Reflective Teacher*. Jossey-Bass: San Francisco, 1995; pp 114-140.
13. Brookfield, S. Critically Reflective Practice. *J. Contin. Educ. Health.* **1998,** *18*(4), 197-205.
14. Glowacki-Dudka, M.; Barnett, N. Connecting Critical Reflection and Group Development in Online Adult Education Classrooms. *Int. J. Teaching and Learning in Higher Education*, **2007,** *19*(1), 43-52.
15. Adams, K. L. (2001). The Critical Incident Questionnaire: A Critical Reflective Teaching Tool, *Exchanges: The Online Journal of Teaching and Learning in the CSU.*
16. Hessler, H. B.; Taggart, A. R. What's Stalling Learning? Using a Formative Assessment Tool to Address Critical Incidents in Class. *Int. J. of Scholarship of Teaching and Learning,* **2011,** *5*(1), 9.
17. Hedberg, P. R. Learning Through Reflective Classroom Practice. *J. Manage. Educ.* **2009,** *33*(1), 10-36.
18. Boyle, J. T.; Nicol, D. J. Using Classroom Communication Systems to Support Interaction and Discussion in Large Class Settings, *Research in Learning Technology*. **2003**, *11*(3).
19. Brookfield, S. D.; Preskill, S. *Discussion as a Way of Teaching: Tools and Techniques for Democratic Classrooms,* Jossey-Bass: San Francisco, 1999.
20. Koretsky, M.; Falconer, J.; Brooks, B.; Gilbuena, D.; Silverstein, D.; Smith, C.; Miletic, M. The AIChE Concept Warehouse: A Web-Based Tool to Promote Conceptual Learning. *Adv. Eng. Educ.* **2014,** *4*(1)

4. DATA ANALYTICS FOR INTERACTIVE VIRTUAL LABORATORIES

Jessie Keeler, Ying Cao, Thomas Ekstedt, & Milo Koretsky

Proceedings of the 2016 American Society for Engineering Education Annual Conference & Exposition

Introduction

We have previously described the development and implementation of a set of *Interactive Virtual Laboratories* (IVLs) in thermodynamics.[1] Each IVL provides a set of activities to address targeted threshold concepts[2] via actively engaging students in a series of actions. The IVLs provide a less abstract and more intuitive access to students by providing a dynamic representation of phenomena at a molecular level. Students are expected to answer numerical questions and, when prompted, predict and explain the effects of macroscopic changes (i.e., pressure, temperature, composition, energy) based on observations of molecular phenomena. Six IVLs are currently available for public use through the *AIChE Concept Warehouse*.

Through this study, we seek to explore ways to use gathered data from student answers to understand learning, supply formative feedback, and provide accountability. The study contains two parts. First, we audio recorded 10 students as they worked through one of two IVLs in an attempt to examine student thinking processes and determine rationale that commonly leads students to submit wrong answers. Second, we implemented the two IVLs in a large, junior level thermodynamics course. All responses were recorded and stored in the *AIChE Concept Warehouse*. In this paper we argue (and will show evidence) that sometimes even though the final answer was not normatively correct, there are still productive thought process behind the answers. One ancillary benefit of this study includes the development and implementation of an automatic grading system for the IVLs. Further improvements that were made to the IVLs are also described.

Background

It is commonly known that many students struggle to learn thermodynamics. Research studies confirm that key concepts of thermodynamics are commonly misunderstood because students develop non-scientific ways of thinking about things such as heat and work before they are introduced to these concepts in a classroom setting.[3-5] For example, van Roon, Sprang, and Verdonk[3] completed a three year study that examined freshmen chemistry students' understanding of heat and work in the context of thermodynamics. Students in this study often identified heat as a state function instead of a path function and routinely tried to conserve it using the First Law.[3] Because thermodynamic concepts are commonly misunderstood, there is a demand for ways to improve student learning. One such method of improvement over traditional lectures is active learning.

Students engage in chemistry and biology laboratories early in their science studies. Specifically, laboratories can promote inquiry-based learning where students engage in meaning making by solving problems, answering questions, or interacting with phenomena instead of didactically being presented with a concept to learn. Numerous studies demonstrate the efficacy of inquiry-based learning in science courses.[6-8] There has also been considerable research that ascertains students learn better with simulations than they do in traditional lectures.[9-11] Stieff and Wilensky[11] implemented computer simulations in undergraduate chemistry courses and interviewed students regarding their interactions with the software. According to their results, "all students took on increasingly more conceptual approaches to solving problems." [9,11]

There is also evidence that students perform better when a virtual laboratory (computer simulation) is used to supplement other learning.[12,13] For example, Zacharia, Olympiou, and Papaevripidou[12] implemented simulations in the form of "virtual manipulatives," in an undergraduate physics course. The participants of the study were randomly assigned into two groups; one group used only physical manipulatives while the other group used physical manipulatives followed by virtual manipulatives. Conceptual tests that were administered before, during, and after the study showed that using both manipulatives in sequence improved the students' conceptual understanding more so than using the physical manipulatives alone.[12] Virtual laboratories are also more flexible than physical laboratories and allow visual representations of phenomena not accessible in a physical laboratory – such as on the molecular scale.

The *Interactive Virtual Laboratories* used in this study were designed to help students master “threshold concepts” in thermodynamics.[1] According to Meyer and Land[2], there are four characteristics of a threshold concept: troublesome, transformative, irreversible, and integrative.[2] Troublesome refers to the difficulty of the concept and the fact that students often struggle with it. Transformative means it alters the way students approach the discipline and related knowledge. A threshold concept is irreversible in the sense that once students correctly understand it they will not return to the more simplistic or uninformed view that they held earlier. Integrative refers to how understanding a threshold concept allows students to make previously unseen connections between aspects of the course.

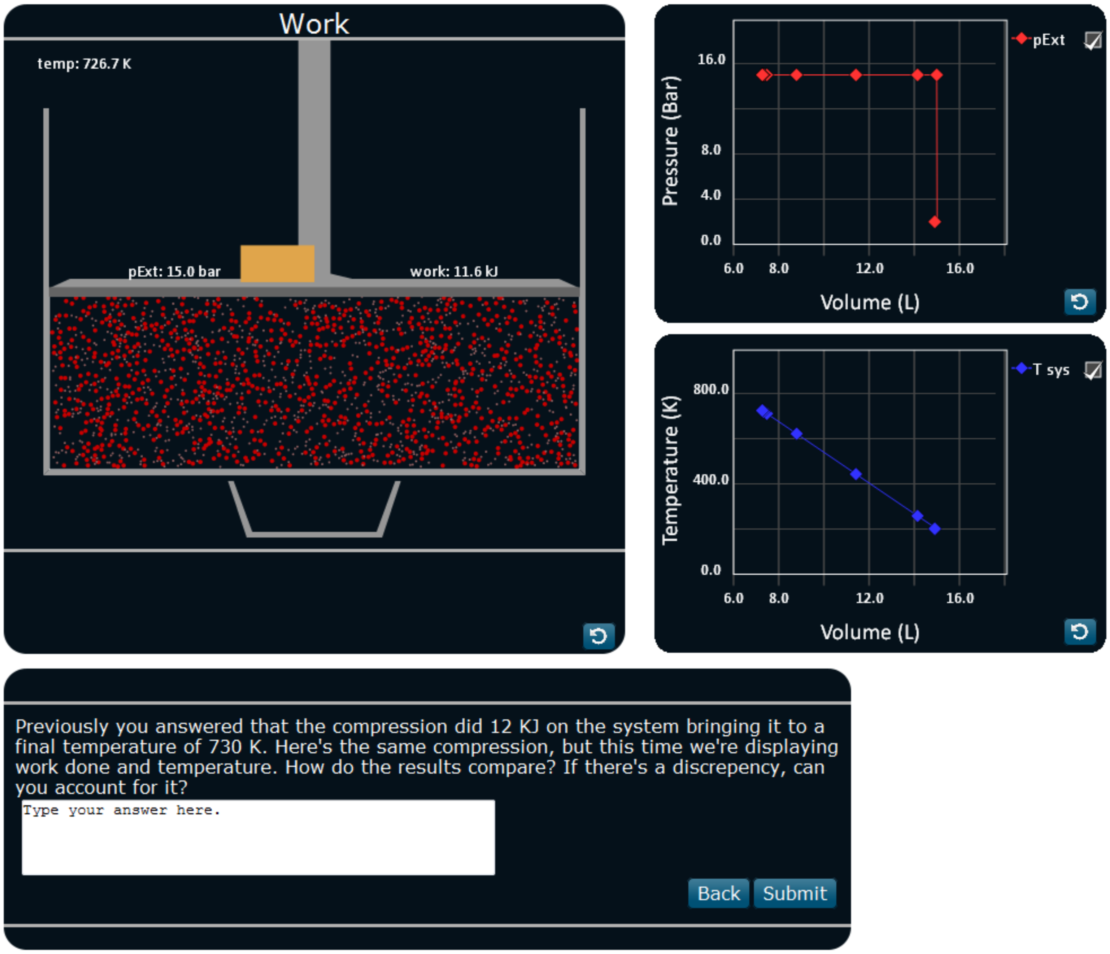
IVL Overview

The IVLs were constructed based on this active learning pedagogy and directed towards undergraduate thermodynamics students.[1] In the IVLs, students are guided through a set of *frames* where they are asked to respond to questions that ask them to predict, calculate, manipulate, observe, or reflect on phenomena related to the specific concept.

Figure 4.1 presents an example frame of the Work IVL, one of the six available IVLs. This frame includes the three main parts of a typical IVL frame, (i) a box containing the molecular simulation that students are asked with manipulating in certain ways - in this IVL, students place a block on a piston and observe as the system of gas molecules is compressed; (ii) a macroscopic, graphical representation of the simulated phenomena (located to the right of the molecular simulation) – the two graphs in Figure 4.1 allow students to check their answers they submitted on previous frames; and (iii) a box to read instructions and provide answers to questions (below the molecular simulation).

The *Interactive Virtual Laboratories* were designed to help students master “threshold concepts” in thermodynamics.[1] Each IVL begins with questions asking students to explain or define the threshold concept that the IVL targets. The students then predict what will happen when the system undergoes a process, such as supplying work. After students complete their calculations they interact with a simulation to see if their predictions are correct. Once students observe the simulation they are asked to explain how the results compare to their predictions. This pattern is progressively repeated until the end where students are prompted to reflect on the content of the lab and asked to define the threshold concept again.

There are a total of six *Interactive Virtual Laboratories* available on the *AIChE Concept Warehouse* and we focused on two IVLs for this study. We used the "Work Simulation" and the "Heat Capacity Simulation." The Work IVL focuses on pressure-volume (Pv) work as an energy transfer process and covers the threshold concept of how pressure-volume work adds energy to a system. The Heat Capacity IVL examines the definition of heat capacity and the threshold concept for this lab is the difference between constant volume heat capacity and constant pressure heat capacity. Previous papers describe the development and implementation of the IVLs in more detail.[1]



Previously you answered that the compression did 12 KJ on the system bringing it to a final temperature of 730 K. Here's the same compression, but this time we're displaying work done and temperature. How do the results compare? If there's a discrepency, can you account for it?

Type your answer here.

Figure 0.1 Sample frame from Work IVL

Methods

Participants

Participants in this study were enrolled in a junior-level engineering course titled "Thermodynamics" at a large, public university in the Northwestern United States. A total of 241 chemical, environmental, and biological engineering students participated. This is a required course for each of the three programs; main topics covered in the course include the first and second laws of thermodynamics, entropy, equations of state, and the thermodynamic web.[14]

Context

The course consisted of two 50 minute, instructor-led lecture sessions on Mondays and Wednesdays as well as two 50 minute, graduate-student-facilitated studio sessions on Tuesdays and Thursdays. There were six different studio sections with an average of 40 students in each section. The lectures were used to introduce content whereas the studios allowed the students to apply what they learned in lecture to practice problems via an inquiry-based worksheet. The Heat Capacity was assigned during studio and therefore students completed it in groups. The Work IVL was assigned as homework and the students completed it individually.

Data collection

Based on the total number of students enrolled in the class and the number of questions in each IVL, slightly more than 7,000 student responses were collected. In addition to the student responses from the general class implementation, 10 students agreed to be audio recorded as they completed the IVL using a think-aloud protocol. Data are only reported for students who agreed to participate and signed an informed consent form approved by the Institutional Review Board. The recordings were transcribed and each student participant was given a pseudonym to protect their identity.

For each of the three assigned IVLs, we collected two sets of responses. One response represented a stronger performing team and the other a weaker performing team; these designations were based on the answers the teams provided to the procedural questions. The stronger performing teams answered more procedural questions correctly than the weaker performing team did. An overview of the recording data we obtained is provided in Table 4.1.

Table 0.1 Recording data overview

IVL	Type of recording	Sets of responses	Total student participants
Work	Individual	2	2
Heat Capacity	Group	2	8

We did not provide specific conceptual prompts to the groups during data collection. Instead, we asked students to think aloud and talk with their group members like they would during a normal studio session. Because we recorded individuals working through the Work IVL, we occasionally asked them to explain why they used certain equations or how they got an answer if they did not elaborate on their own. Essentially, we played the role of a confused group member to better understand their thinking processes when they were not clear or explicit.

Data analysis

As part of our analysis on student answers to questions in the IVLs, we coded each question in the lab into one of four categorizes according to what type of thinking each question elicited: procedural, conceptual, prediction, and reflection.[14] Table 4.2 contains a brief explanation of the codes.

Table 0.2 IVL question coding descriptions

Thinking Elicited	
Procedural	Elicit computation or numerical graphical interpretation. The answers are typically numerical.
Conceptual	Elicit students’ conceptual interpretation of data or information in order to explain complex phenomena. The answers are typically in text.
Prediction	Elicit students’ anticipation about what will happen if they make a change to the system. The answers can be a mixture of text and number.
Reflection	Elicit student thinking back to previous problems and comparing how results differ due to changes. The answers can be a mixture of text and number.

The codes allowed us to get an overview of what types of questions each lab asked. It also allowed for an easier comparison and analysis on student responses to a group of questions of a same type. Table 4.3 details the coding breakdown and total question count of each IVL.

Table 0.3 IVL question coding count

	Work	**Heat Capacity**
Procedural	9	3
Conceptual	6	4
Reflection	3	5
Prediction	0	1
Question Total	18	13

We examined the transcripts of the audio recordings and identified the student thinking processes that led them to the answers they input in the IVLs. We especially tried to recognize (1) common misconceptions that led students to common wrong numerical answers of procedural questions; (2) productive discussion in conceptual questions regardless of whether answers to procedural questions were accurate; and (3) reasoning that was canonical and also led students to arrive at the correct answers to procedural questions. For each incorrect procedural question that the groups or individuals submitted, we determined how the answer was reached and whether or not it was the result of a misconception or a calculation error. This analysis provided us a spectrum of student thinking and responses, in continuum, from wrong-answers with wrong-reasoning, to partially-correct reasoning, to correct-answers with correct-reasoning.

Results

In this section we present cases of students working through the Work and Heat Capacity IVLs. For each case, we include a detailed comparison of the student thinking processes as the individuals, or groups, worked through both a procedural and a conceptual question.

Work: Conceptual Questions

We recorded two individuals as they completed the Work IVL. An example of a conceptual question from the Work IVL is shown in Figure 4.2. The figure shows the third frame (out of total 18 frames) in the IVL.

Before this conceptual question students had answered two multiple choice conceptual questions. The first question asked the students to select an equation that best represented the work done in an adiabatic system. It included a problem statement with information about the system, 4 possible answer choices, and a diagram of the system to the right of the question. The correct answer choice states, in a symbolic format, that work (W) equals the product of minus external pressure (-P) and change in volume (ΔV). The question on the second frame presented students with several different versions of the First Law of Thermodynamics and asked students to choose which one correctly described the adiabatic system. The question contained a problem statement with relevant equations, 4 possible answer choices, and the same diagram as the first frame. The correct answer choice was $n\ c_v\ \Delta T=-P(\text{external})\Delta V$, which indicates the equivalence between the change in internal energy and the work done (because no heat was transferred). Students could not progress to the third frame until they correctly answered both multiple choice questions in the first two frames. Both students (Gerry and Lucy) answered the questions correctly and moved on with Frame 3.

$$nc_v\Delta T = -P_{ext}\Delta V$$

From the equation above we see that temperature increases as we do work by decreasing volume. Temperature is an expression is molecular kinetic energy, so as the system is compressed, the molecules must speed up. These ideal gas molecules can be thought of as perfectly elastic bouncy balls. Using the movable wall above, can you determine what event causes the molecule's speed to change? Can you explain why that would cause a temperature change in many molecules?

Figure 0.2 Conceptual question from Work IVL

The question on Frame 3 (shown in Figure 4.2) continued with the question students had chosen in Frame 2. Students were asked to manipulate a horizontal movable wall on top of the container and observe the change in the speed of a gas molecule in the container. They were then prompted to explain why this event would cause a temperature change in a system with numerous molecules. When responding to the question shown in Figure 4.2, Gerry thought out loud and made clear connections between what she observed in the simulation and the information included in the problem statement:

> **Gerry:** I think by changing the volume I can see the speed is changing in the simulation. Which means molecules' kinetic energy is changing and from this problem kinetic energy change will lead to temperature change so it makes sense that delta V is related to delta T.

Gerry then typed her answer to the question on the frame: "Volume change leads molecule's speed changes which leads temperature change." Gerry did not relate the increase in molecular speed to the molecule colliding with the moving wall; instead, she describes the overall volume change as the reason the kinetic energy increases. It appears that Gerry based her answer on observation but did not explicitly infer the cause, unlike Lucy.

In the audio recording, student Lucy talked about the idea of transferring momentum to the molecule from the moving wall:

> **Lucy:** Um, temperature is a measure of average kinetic energy… by transferring momentum from the wall to the molecules you speed them up, causing the temperature of system to increase.

Lucy was typing into the answer box while she was talking: "A transfer in momentum from the wall to the 'molecule.' Temp is a measure of average kinetic energy. By transferring momentum from the wall to the molecules, you speed them up causing the temp of the sys to increase."

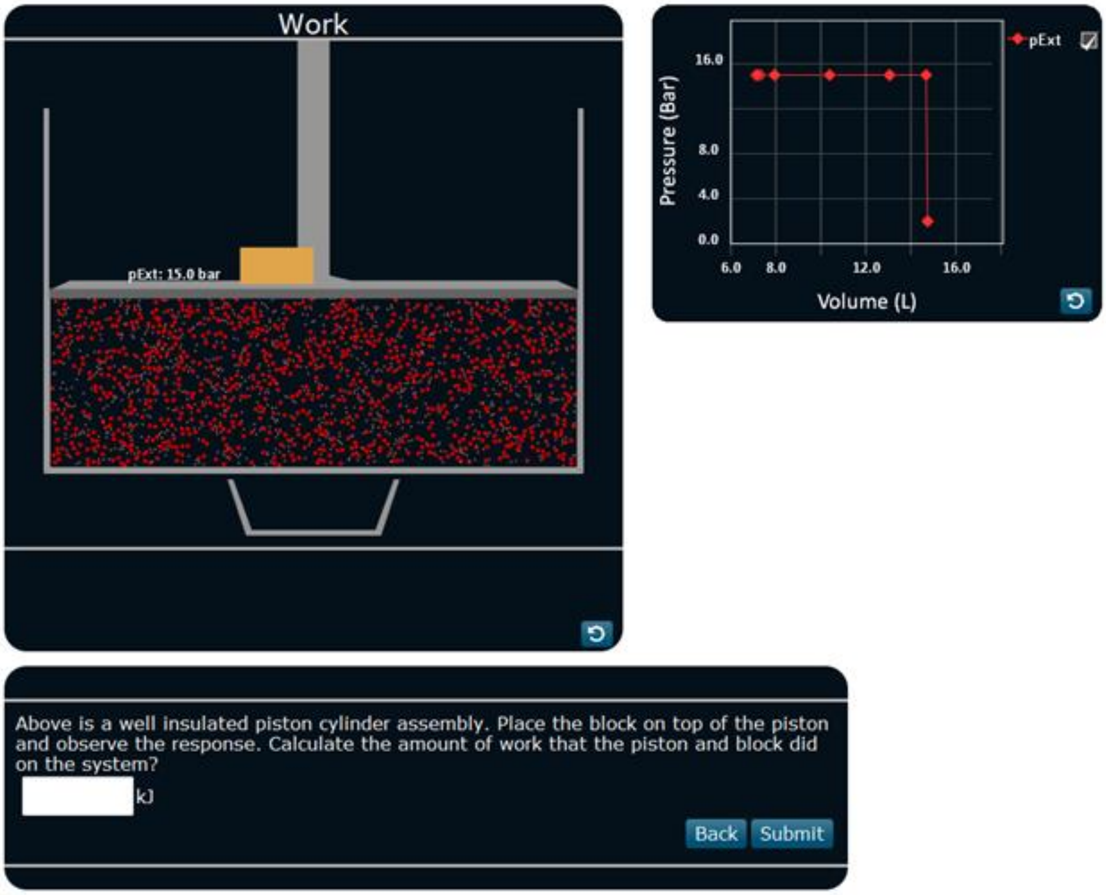
Both students linked an increase in the molecule's speed to an increase in its kinetic energy, which would result in an increase in temperature. Lucy's response was consistent with the canonical explanation and both her reasoning and answer were considered correct when we coded them. Gerry did not directly associate the increased speed of the molecule to the collision between the molecule and the wall; therefore, we considered her reasoning and answer as partially-correct.

Work: Procedural Questions

After the students completed the conceptual question shown in Frame 3, they ran a simulation in which they virtually dropped a block onto a piston and observed the piston compress a system of many molecules. They answered two conceptual questions in Frame 5 relating to this simulation before they reached their first procedural question in Frame 6 (Frame 4 has no question. It is the explanation of previous simulation).

For the question in Frame 6 (shown in Figure 4.3), students were also given a graph that displayed the changes in volume and pressure as the system is compressed. Students were prompted to calculate how much work the piston and the block did on the system. Both students recognized they needed to use the information in the graph to the right of the simulation. However, two different errors caused Gerry to submit the wrong answer:

> **Gerry:** This is insulated which means Q is zero. So, 'calculate the amount of work that the piston and block did on the system'. So as I can see in PV graph, this is [an] irreversible process, so the area under the PV curve will be the work that the piston did on the system. Since [the] system got the work, work will be a positive number. So I can just simply get the area under the curve which would be, I don't know, 15 times 15? So 225 approximately.



Above is a well insulated piston cylinder assembly. Place the block on top of the piston and observe the response. Calculate the amount of work that the piston and block did on the system?

Figure 0.3 Procedural question from Work 1

Gerry was correct that it is an irreversible process and that work is equal to pressure times the change in volume, but when she was getting these values from the graph she neglected to recognize that the volume axis did not start at zero. Therefore, she used the wrong value for her volume change. She also did not pay attention to units when she calculated for work.

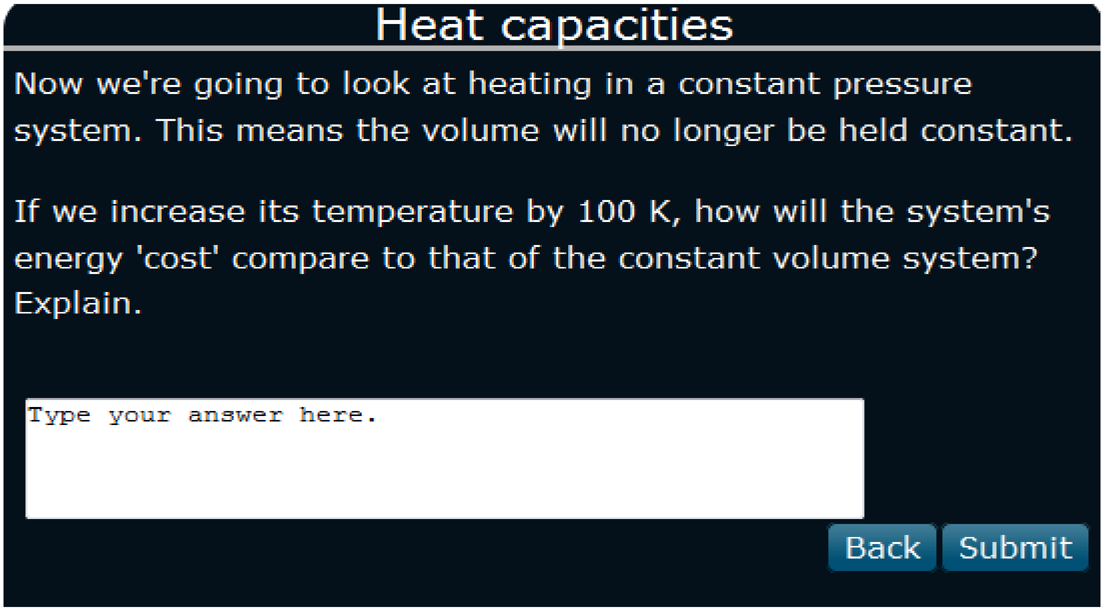
By reading the graph more closely and converting her units, Lucy submitted the correct answer:

> **Lucy:** So I at first, I know that work is pressure times the change in volume and so on this little graph over here I see just that, except it's confusing because I don't know the volume change so I have to look at this graph to observe the volume change… All right, so about negative—minus 8 liters, so that's the change in volume I got, so that divided by 1,000 gives me meters cubed…15 bar, 15 e to the 5th, times 0.008, and it was negative, too, but then negative and a negative makes a positive because energy is being done on me, the system, divided by 1,000 to get me in kilojoules, because I think you're going to ask for kilojoules. I hope this is right but it could be wrong. I'm going to say it's around 15- 12, I'm going to say it's around 12.

The individuals that worked through the Work IVL represent two different cases; responses from Lucy represent correct answers with correct reasoning for both the conceptual and procedural questions, whereas responses from Gerry represent partially-correct answer and partially-correct reasoning for the conceptual question and wrong answer with partially-correct reasoning for the procedural question. Although Gerry used the appropriate equation in her procedural calculation, she did not account for her units nor did she correctly read the graph when gathering values to use in her equation. This mistake represents procedural errors instead of a misconception.

Heat Capacity: Conceptual Questions

For the Heat Capacity IVL, Group 1 was the weaker performing group and Group 2 was the stronger performing group according to the scores they received in the on procedural questions. Although Group 1 answered the majority of the procedural questions inaccurately, they still answered the conceptual questions correctly. A conceptual question that highlights the threshold concept of the Heat Capacity IVL is shown in Figure 4.4.



Heat capacities

Now we're going to look at heating in a constant pressure system. This means the volume will no longer be held constant.

If we increase its temperature by 100 K, how will the system's energy 'cost' compare to that of the constant volume system? Explain.

Figure 0.4 Conceptual question from Heat Capacity IVL

Figure 4.4 shows the screenshot of Frame 4 out of 10 total frames in the Heat Capacity IVL. In Frame 4 students were asked a conceptual question about *constant pressure* heat capacity, c_P. Before students reached this question they worked through previous frames about a *constant volume* system. In the first frame they were asked the question "Using your own words, how would you explain the concept of a heat capacity to a high school senior?" In the second frame they were asked “How much energy should heating this system by 100 K 'cost'?” The system in question contained 0.5 moles of an ideal, monatomic gas of a fixed volume. The third frame contained a simulation where they could pull a slider that simulates the switch of a heater to virtually “heat” a container of gas of a fixed volume (molecules represented by dots moving and bouncing around in the container) until the number on top of the frame that indicates the temperature achieved a 100 degree increase in temperature. This simulation allowed students to test their prediction from the second frame. They were instructed to observe how much energy was being used as they heated the system (the amount of energy used was also shown on top of the frame) and then describe how the total required amount they observed in Frame 3 compared to their predicted value in Frame 2 (more information about the frames are available on *AIChE Concept Warehouse* website).

In Frame 4 that follows, the question (shown in Figure 4.4) asks students to compare the amount of energy required to increase the temperature of a *constant pressure* system versus the energy required to increase the same degrees of temperature of a *constant volume* system (with which students had worked during the first three frames). Group 1 and Group 2 reached the same conclusion through two different ways of reasoning.

Below is the transcript of student discussion when Group 1 was working on this frame. April and Ben represent two students in this group.

> **April:** So c_p equals c_v plus R. So if we're keeping pressure constant instead of volume then we should require more energy. We need—so we're going to have more energy.
> **Ben:** That makes sense because you're probably also going to do work on the system.
> **April:** Yeah. So, we're putting that, we're going to put: the energy cost is going to be higher because the system has to do work to expand, um, to make the volume expand.

This excerpt shows that April started with the (correct) relationship between the heat capacity of a constant pressure system and the heat capacity of a constant volume system ($c_p = c_v + R$) and determined how the energy requirement varied. This excerpt is also an example of co-construction and shows the benefits of group work. Ben added his understanding of the relationship between the energy cost and the work done (more energy required because work was involved). April, when inputting their response, added the relationship between work and volume (the system has to do work to make the volume expand). April submitted the following written answer: "the energy cost is going to be higher because the system has to do work to make the volume expand." The response that Ben submitted was: "The energy will be higher due to the system requiring to increase the volume." The third student in Group 1, Donna, provided the following written response: "The system's energy cost should be higher than constant volume because the system has to do work to expand." All these showed the students' understandings of the concepts in thermodynamics related to the simulation. It is also interesting to note that it was Ben who brought up the notion of work, but he was the only one in this group who did not include "work" in his submitted answer.

Group 2 discussed the differences of energy requirement between the constant volume system and the constant pressure system in terms of the First Law, although they did not explicitly name it. Below is the transcript of the discussion of Group 2 when they were working on Frame 4. Leslie, Ron, and Chris represent the students in this group.

> **Leslie:** Okay, because for the previous—when we had a constant volume no work was being done. So you just had delta U equals Q. And now you have a –W, or negative work, like overall net work would be negative. Okay, I see what you're saying, because you were saying the energy cost would be higher. Internal energy is being kept constant, because for an ideal gas it only depends on temperature and we're increasing it a certain amount of temperature, right?
> **Ron:** The change in internal energy is the same here as in the last problem, you're saying?
> **Leslie:** Mhmm. And then it's just a matter of how much—
> **Chris:** How much heat you have to apply to get—
> **Leslie:** To get it to the same temperature

The conversation shows that the students know the First Law of Thermodynamics and can relate changes of internal energy (delta U), to heat (Q), and work (W). They also identify the relationship between temperature and internal energy for an ideal gas (that the internal energy of an ideal gas only depends on temperature) and the relationship between the gas's volume and the work done (that if the volume expands, the system has done a negative work). This thought process set the students up to correctly answer the procedural questions they encountered later in the IVL. The actual submitted answers from students in this group are shown in Table 4.4. The responses were basically consistent with their discussion.

Table 0.4 Group 2 student responses to the conceptual question in Heat Capacity IVL.

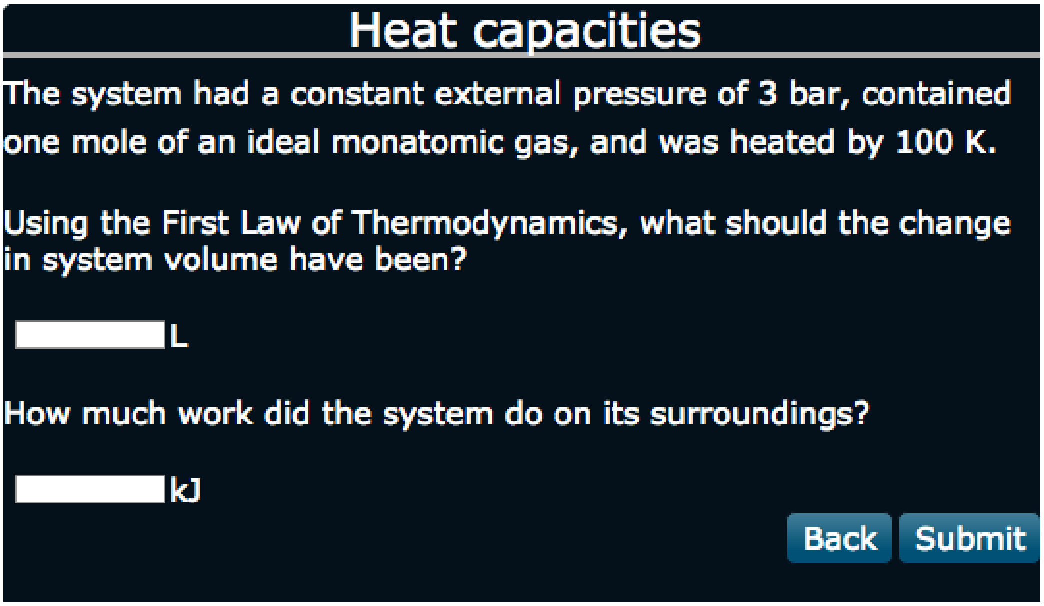
Student	Submitted response
Chris	The energy cost is higher in the constant pressure case because the internal energy given to the system via heat is lost through work. Thus, it will take a longer time to increase the temperature requiring more energy.
Leslie	The energy cost would be higher (i.e. more energy would need to be put into the system in order to reach the same temperature). If the energy cost is being defined as the heat put into the system, it would cost more, because in a constant pressure system the volume would change (expand in this case) and thus work would be done. By doing work, more heat needs to be put in to acheive the same change in internal energy
Ron	I am defining energy cost as heat transferred to the system in order to raise the temperature 100K. In a constant pressure system, the energy cost will be greater than it was in a constant volume system because the system will be doing work on the surroundings throughout the heating process. The work done results in a smaller increase in internal energy per unit of heat transferred.

Tammy	Its an expansion problem, Work is negative, Q is positive, Using nCpdeltaT to calculate work. Constant Volume no work is being done. Using equation deltaU=nCvdeltaT
Ann	The constant pressure system would be more costly. This is because when energy is added to the system the volume expands but this means work has to move the walls of the system. This work energy leaving the system means that for every unit of energy entering the system, some portion will leave as work. In the constant volume situation, no work energy left the system.

Group 1 approached the energy requirement question in Frame 4 by observing the relationship between c_v, or heat capacity of a constant volume system, and c_p, heat capacity of a constant pressure system. They knew that for an ideal monatomic gas $c_p = c_v + R$. Therefore, they determined that because c_p is larger than c_v more energy would be required in the scenario with the constant pressure system. They explained this increased energy requirement from a physical perspective; they thought about what would happen to the system if the volume wasn't held constant and the temperature were increased. They determined that the system would expand and thus work would be done. It is clear from their written responses that the students in Group 2, with the possible exception of Tammy (who seemed to reason through the conceptual question in a procedural sense via the application of equations), also have a deep conceptual understanding of the different energy requirements between a constant pressure system and a constant volume system.

Heat Capacity: Procedural Questions

Following the conceptual question in Frame 4 (shown in Figure 4.4) Frame 5 asked students to reflect what they have done in Frames 1 through 4. Frame 6 then contained two procedural questions as shown in Figure 4.5.



Heat capacities

The system had a constant external pressure of 3 bar, contained one mole of an ideal monatomic gas, and was heated by 100 K.

Using the First Law of Thermodynamics, what should the change in system volume have been?

☐ L

How much work did the system do on its surroundings?

☐ kJ

Figure 0.5 Procedural Questions from the Heat Capacity IVL

Group 1 had misconceptions related to the First Law that caused them to submit wrong numerical answers in Frame 6. Group 1 incorrectly set up the equation to calculate for the change in volume but used the correct equation to solve for the amount of work done on the surroundings. However, because their numbers from the first question were incorrect they did not achieve the correct answer for the second question either.

To solve for the first problem, Group 1 used the equation Q=W, or heat added to the system equals work done by the system.

> **Donna:** Okay, so then we can just solve for delta V. c_p delta T over P external, does that look right?
> **April:** Because it's just Q=W, that's just W and we want to calculate that. So how much work was done on the surroundings is just that [previous answer of ΔV] times 3 [the value of external pressure]…

They claimed all of the energy added to the system by heating it up was transferred to the surroundings in the form of work as the container expanded. This is a misconception; they did not account for the change of internal energy in their equation. Internal energy is a function of temperature, and thus, the change in internal energy was not zero for this process because it was not isothermal. The change in internal energy should be the same for both the constant volume and constant pressure systems because both were experiencing the same change in temperature.

However, more energy needed to be added to the constant pressure system to achieve the desired temperature change because as the system expanded part of the energy was being lost to the surroundings in the form of work.

Because students in Group 1 did not include the change in internal energy in their equation, they did not get the correct answer for the first question (the change in volume). However, they still had some correct understanding of the relationship between heat, heat capacity, and change in temperature (that is, $Q = c_p \Delta T$). In answering the second question in this frame they also used the correct relationship between work and change in volume (that is, $W = P \Delta V$).

Group 2 correctly solved for the procedural questions displayed in Figure 4.5 by introducing enthalpy into their calculation.

> **Chris:** I used delta U equals n c_v delta T. Then you know delta H is equal to U plus PV and since you know delta U and you can find delta H since it's an ideal monatomic gas you have c_v and you can find c_p from c_v.
> **Tammy:** So you have delta U n C delta T and then H equals U plus PV?
> **Chris:** Yeah.
> **Tammy:** And then PV…
> **Chris:** You can solve for that. Because you know c_p is equal to R plus c_v
> **Tammy:** Oh, this should be c_p [delta U]
> **Chris:** No, that's c_v for delta U. And then c_p for delta H
> **Tammy:** Oh, okay, the same equation but c_p
> **Chris:** Yeah
> **Tammy:** And then you can solve for PV

Students in Group 2 used the definition of enthalpy, which states for a constant pressure system $H=U + PV$. They reasoned that the change in enthalpy was equal to the change in internal energy plus the product of pressure and (the change in) volume. They then stated that $\Delta U=n\ c_v\ \Delta T$ and $\Delta H=n\ c_p\ \Delta T$. After making these substitutions and rearranging Group 2 was able to correctly solve for the change in volume.

The groups that worked through the Heat Capacity IVL represent two different cases; responses from Group 2 represent correct answers with correct reasoning for both the conceptual and procedural questions, whereas responses from Group 1 represent correct answer and correct reasoning for the conceptual question and wrong answers and partially wrong reasoning for the procedural questions. Group 1 seemed to relate heat capacity to heat rather than a change in internal energy.

Discussion

Through this study we were able to examine student thinking processes as they work through these simulations. We also discovered errors and/or misconceptions some students have regarding thermodynamic principles. We were easily able to determine how many students might have shared the errors and/or misconception by looking at their numerical answers.

Using the information gained from examining the Work IVL recordings and looking at the bigger scale of student answers of 241 total students, we were able to identify students that made calculation errors similar to Gerry when they responded to the procedural question shown in Figure 4.3. Based on the student answers submitted for the Work IVL, 32 of the total 240 students who participated made a unit error in their calculations and 7 of those 240 students read the PV graph incorrectly. Thus, approximately 16% of the students who answered incorrectly did so most likely because of calculation errors as opposed to misconceptions.

By identifying and detailing the misconceptions Group 1 had in the Heat Capacity IVL it helped us to identify students with analogous misunderstandings in the future. For example, by examining the answers students submitted for the Heat Capacity IVL, we discovered that 24 of the total 237 students answered "6.93" for the change in volume We speculated with confidence that these other 19 students solved for the change in volume by using the same equation as the 3 students in Group 1: Q=W. Therefore, from this information we can assume that approximately 10% of the students have misconceptions involving internal energy.

Working toward the automatic grading system as one of our ultimate products of this project, we wanted to understand the thinking processes that led students to the answers they put in the boxes in the IVL. We have found that even if students' final answers were not accurate, there are still productive ideas behind the answers. Therefore, the next challenge is to accommodate these thought processes into the grading system.

IVL Improvements

Besides examining student thinking process and misconceptions, another related goal of this project was to inform instructors about student thinking and performance. The project allows instructors to gather data from the IVL labs and use that information to better inform teaching and thus improve learning. To fulfill this objective we developed an automatic grading system via an algorithm based off of student responses to questions in the IVLs. This grading mechanism will serve as a measure of student performance and allow the instructor to see specific areas where students are struggling. Because students will have access to this paper we will not detail how the questions are scored. The grading mechanism is currently being integrated into the *AIChE Concept Warehouse*.

We are also working on implementing a graphical display of student scores that will be visible to the instructor. This will provide the instructor with a quick overview of student performance for each IVL; without having to scroll through individual scores, the instructor will have an idea of whether or not students understood the concepts covered in each IVL based on a distribution of student scores. Based on the student answers we recorded during this study, Figure 4.6 shows a mock-up of the instructor display.

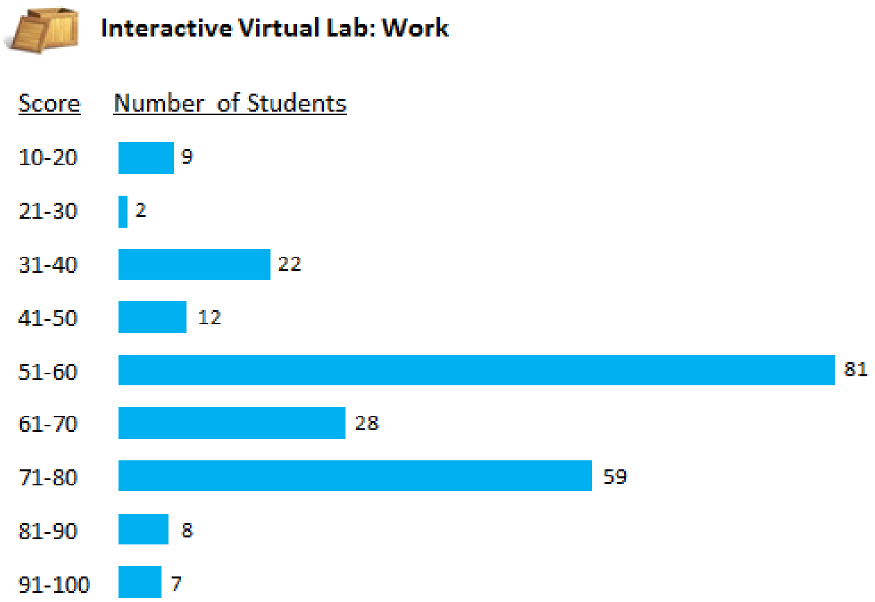
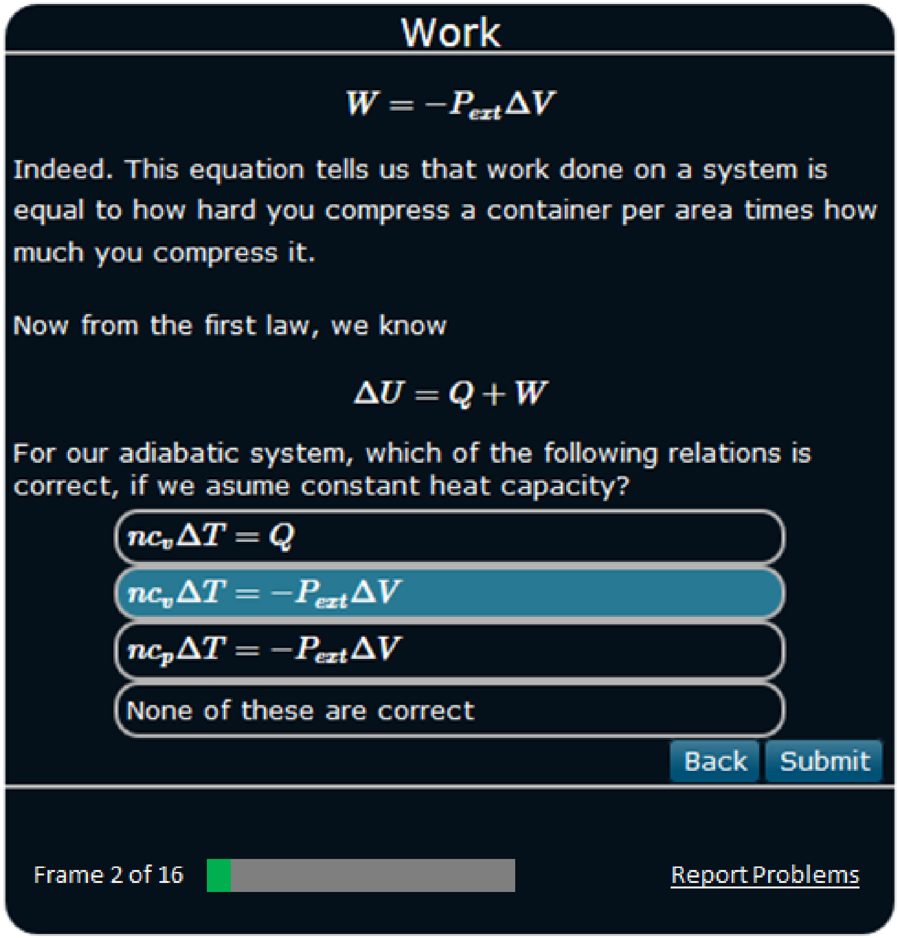


Figure 0.6 Mock-up of graphical distribution of student scores on the Work IVL (from class data)

Students working through the IVLs often commented on how it would be helpful if they had a measurement of how far into the simulation they were. To remedy this we are in the process of numbering each frame of the IVLs as well as displayed the total number of frames. Not only will this allow students to gauge their process but it also makes referencing frames in the IVLs more universal. For example, if a student is having a technical difficulty with a simulation they can now clearly articulate which frame they were on when the problem occurred. We are also in the process of implementing an interface that will allow students to submit any problems they had while completing the IVL. For example, if a student completes an entire IVL but when they hit submit they are directed back to the beginning and their answers are not saved they can submit an explanation of what occurred without their score being affected. The submitted problems will be available to the facilitator in charge of recording the scores as well as someone in charge of fixing IVL-related technical problems. Based on these improvements, Figure 4.7 shows a mock up with these features.



Work

$$W = -P_{ext}\Delta V$$

Indeed. This equation tells us that work done on a system is equal to how hard you compress a container per area times how much you compress it.

Now from the first law, we know

$$\Delta U = Q + W$$

For our adiabatic system, which of the following relations is correct, if we asume constant heat capacity?

- $nc_v\Delta T = Q$
- $nc_v\Delta T = -P_{ext}\Delta V$
- $nc_p\Delta T = -P_{ext}\Delta V$
- None of these are correct

Figure 0.7 Mock-up of progress display and link to submit problems (bottom)

We are currently creating a guide for IVL usage. This guide will function as a brief list of suggestions to help students run the simulations as smoothly as possible. For example, the guide will contain browser and device suggestions, such as no Internet Explorer and laptops are preferred. This guide will most likely be integrated into the *AIChE Concept Warehouse* where both instructors and students can access it. We will encourage instructors to advise students to read the guide before they use any of the IVLs. By creating this guide we hope to limit the amount of preventable errors that occur with the IVLs, such as students trying to use tablets and not being able to fully interact with the simulations.

References

1. Koretsky, M. D. (2014). Development and implementation of interactive virtual laboratories to help students learn threshold concepts in thermodynamics --Year 1. Proceedings from ASEE 2014: American Society of Engineering Education Annual Conference. Indianapolis, IN.
2. Meyer, J. H. F., & R. Land. (2003). Enhancing teaching-learning environments in undergraduate courses. Occasional Report, Centre for Teaching, Learning and Assessment, The University of Edinburgh
3. van Roon, P. H., van Sprang, H. F., & Verdonk, A. H. (1994). 'Work' and 'heat': One a road to thermodynamics, *International Journal Science Education*, 16(2), 131-144.
4. Nilsson, T., & Niedderer, H. (2012). An analytical tool to determine students' use of volume and pressure when describing expansion work and technical work. *Chemistry Education Research and Practice*, 13, 348-356.
5. Bain, K., Moon, A., Mack, M. R., & Towns, M. H. (2014). A review of research on the teaching and learning of thermodynamics at the university level. *Chemistry Education Research and Practice*, DOI: 10.1039/C4RP00011K
6. Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99-107.
7. Furtak, E. M., Seidel T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research*, *82,* 300-329.
8. Lewis, S. E., & Lewis, J. E. (2005). Departing from lectures: An evaluation of peer-led guided inquiry alternative. *Journal of Chemistry Education, 82*, 135-139.
9. Smetana, L. K., Bell, R. L. (2012) Computer simulations to support science instruction and learning: A critical review of the literature. *International Journal of Science Education, 34*(9), 1337-1370
10. Scalise, K., Timms, M., Moorjani, A., Clark, L., Holtermann, K., & Irvin, P. S. (2011). Student learning in science simulations: Design features that promote learning gains. *Journal of Research in Science Teaching, 48*(9), 1050-1078.
11. Stieff, M., & Wilensky, U. (2003). Connected chemistry: Incorporating interactive simulations into the chemistry classroom. *Journal of Science Education and Technology, 12*(3), 285–302.
12. Zacharia, Z. C., Olympiou, G., Papaevripidou. (2008). Effects of experimenting with physical and virtual manipulatives on students' conceptual understanding in heat and temperature. *Journal of Research in Science Teaching, 45*(9), 1021-1035.

13. Wu, H., Krajcik, J., & Soloway, E. (2001). Promoting understanding of chemical representations: Students’ use of a visualization tool in the classroom. *Journal of Research in Science Teaching, 38*(7), 821–842.
14. Koretsky, M. D. (2013). *Engineering and chemical thermodynamics*. John Wiley & Sons: Hoboken, NJ.
15. Bowen, A. S. (2014). The development and implementation of interactive virtual laboratories. (Honors Thesis). Oregon State University, Corvallis, Oregon.

5. A COMPARATIVE CASE STUDY OF THE ENACTMENT OF ACTIVE LEARNING IN TWO LARGE-ENROLLMENT UNDERGRADUATE SCIENCE COURSES

Jessie Keeler, John Ivanovitch, Milo Koretsky, & Jana Bouwma-Gearhart

Introduction

Although a considerable amount of studies focus on the effectiveness of active learning applied in various scenarios or in different courses, there is still uncertainty surrounding the idea of active learning in the sense that it is such a broad term. What does an instructor mean when they claim to teach an "active learning class"? What do they hope to achieve by structuring their courses like this? What tools will they use to reach these goals and how do said tools tie into their courses? We aim to take a closer look at the design of content in the context of an active learning environment. We want to examine how this content asks students to process the course material and make meaning. Although many comparative studies of active learning environments versus non-active learning environments exist, lacking is a thick description of what active learning is.

While reports on the implementation of active learning tools are common in the research literature, less attention has been given to how actual course content is integrated into these tools, or even how such tools fit together in relation to other important aspects of delivering a class. Also underrepresented in the research literature are detailed characterizations of students' intended thinking processes and meaning-making as they engage in active learning tasks. In this article, we compare the use of two common active teaching tools– audience response systems (ARS) and inquiry-based worksheets (IBWs)–in two different STEM courses (one in biology and the other in bioengineering), with the goals of comparing instructional decision making and types of thinking instructors intend to elicit from students as well as elucidating how course content was integrated into these tools. Overall we found that while the active teaching tools in the two courses appear similar upon first inspection, the way in which students were being asked to engage in problem solving, or demonstrate understanding, varied considerably. For example, in the biology course ARS questions were used primarily to "check in" with students to see if they were correctly interpreting worksheet content (e.g., graphs and models) or to ask students to recall material recently introduced. In the engineering course ARS questions asked students to apply learned concepts to new scenarios towards improving students' conceptual understanding.

This comparative case examines the active teaching practices of two experienced and highly regarded instructors. Both incorporate the same active teaching tools (ARS and IBW) regularly. Through analysis of the similarities and differences in their practices, we seek a more detailed and nuanced understanding of what active learning is and how it can be integrated into a classroom learning environment. Our goal is not to compare learning gains in order to claim one method or implementation strategy is better than another. Rather, we plan to focus on instructor reasoning and instructional decision making. Through examination of the design differences using the same tools, we seek to provide instructors a deeper understanding of the choices they have in implementing active learning and to provide researchers a more precise set of factors to consider in researching instructional design in active learning settings.

Background

Changes in postsecondary educators' teaching to align with instructional practices grounded in research on how students best learn are considered necessary to meeting the demand for career-ready science, technology, engineering and mathematics (STEM) graduates.[1,2] Such changes are particularly relevant in light of empirical evidence supporting claims that the use of these evidence-based instructional practices (EBIPs) promotes student persistence within STEM majors,[3-6] and reduces the achievement gap for student populations traditionally underrepresented in higher education,[7] both of which are seen as fundamental to addressing national decrees for increased numbers of STEM graduates.[8] Additionally, the value that a STEM literate citizenry holds towards addressing or confounding multiple pressing public health and environmental problems (e.g., antibiotic resistant bacteria, decaying infrastructure and global climate change) further highlights the need for better learning and understanding of these subjects, at all grade levels.[9]

Our research focuses on educators' use of teaching tools designed to engage students in the learning process through participation in active learning tasks. *Active learning* has been broadly characterized as anything that students do in the classroom other than passively listening.[10] While this definition is oversimplified, it has been argued that "encouraging the majority of STEM faculty who only lecture to use any form of active instruction (not necessarily the optimal forms of these instructional approaches) may lead to greater gains in student learning productivity".[11] A more comprehensive picture of active learning describes active teaching tools as "instructional practices that engage students in the learning process" and that "require students to do meaningful learning activities and think about what they are doing",[12,13] with a common goal of shifting the paradigm of teaching from educator-centered to learner-centered.[14] For example, in active learning environments students might participate in reflective writings, problem solving, or inquiry-based worksheets, individually or collaboratively. While active teaching tools engage students in the learning process they do so at different levels, ranging from "not passive" to active, towards constructive and interactive– the hypothesis being, the more engaged a student is the more they learn and the better they understand the material.[15] Thus, teaching for active learning is more complex than moving students beyond passively listening.

Many researchers have articulated the need for active learning to become more common place within their respective disciplines. Within the biology community, Wood[16] expressed the necessity of practices such as group work, formative assessment, and active engagement (among other innovations) in undergraduate courses. He argued that group work shifts the classroom environment from competitive to collaborative, formative assessment supplies immediate feedback to both students and instructors related to how well certain topics are understood as opposed to waiting for homework assignments and exams to be graded, and participating in active engagement as opposed to passive lectures increases student learning gains.[16] To support these claims, lot of research has focused on the benefits of active learning. For example, Smith, Wood, Krauter, and Knight[17] performed a study examining the effectiveness of peer discussion and instructor explanation after using ARS questions in class. They compared learning gains from three different groups of students: those who answered ARS questions then participated in peer discussions, those who answered ARS questions then listened to instructor explanations, and those who answered ARS question and then participated in peer discussions followed by instructor explanations. They found that although all students learn from the concept-questions, but the highest learning gains were seen in the combination peer discussion and instructor explanation group.[17] Similarly, Armbruster et al.[18] implemented active and problem-based learning strategies into an undergraduate biology lecture while they simultaneously shifted the course towards a more student-centered learning environment. After comparing versions of the course with and without the added active learning elements, they determined that not only did student attitudes improve with the new instructional design, but student scores on final exams also increased.[18]

A potential limitation of generalizing active instruction at large scales is that course content and intended learner outcomes are diverse, not only across departments, but also between universities and educators, making comparisons of active learning in these varied environments problematic. Further, instructors may claim they are implementing active teaching tools, but do so only at the surface-level, or fail to enact such tools as originally designed.[19] For instance, Borrego and colleagues[20] examined the fidelity of implementation of 16 research-based active learning tools, such as peer instruction and collaborative learning, by surveying approximately 400 instructors in engineering science courses. These researchers described fidelity of implementation as how closely an implemented instructional practice matched the intentions of those who developed it.[20] Depending on how involved the requirements were for an individual strategy to be implemented correctly, fidelity ranged from 80% to as low as 11% for 11 of the 16 strategies used in the study.[20] As such, the field requires more detailed inquiry into how active teaching tools implemented across different disciplines and institutions vary, and insights into why.

Methods

Described in greater detail below, data sources for our study into the enactment of active teaching tools consisted of semi-structured interviews with course instructors and class artifacts such as inquiry-based worksheets, course syllabi, ARS questions and student responses to these questions. Research questions guiding our investigation included, (i) what types of thinking processes and meaning making are educators intending to elicit from students during the use of ARS questions and IBWs and (ii) what are the similarities and differences between the enactment of these strategies in the two courses studied?

Case Selection

Our selection of cases was based on shared course elements: (i) use of the same active teaching tools (ARS and IBWs), (ii) large student enrollments and (iii) required courses for students majoring in the respective disciplines. A further consideration was that both instructors had continuously engaged in these teaching tools for several terms prior to our investigation and were rated by peers and students as excellent teachers. We felt it important to select cases where the instructors were familiar with the enactment of active learning tools in their classes to better ensure implementation in the course was stable; an inexperienced instructor may not have a consistent enactment of these tools. Below, we provide a brief overview of the two courses comprising our comparative case study:

Case 1: The Junior-level biology course, Advanced Anatomy and Physiology, was delivered to approximately 165 students majoring in biology. One class period each week was committed to inquiry-based worksheet sessions, punctuated by ARS questions and mini-lectures. These sessions were instructor led, but also had seven trained undergraduate learning assistants to facilitate group work.

Case 2: The sophomore-level engineering course, Material Balances, was delivered to approximately 300 biological, chemical, and environmental engineering students. One day a week ARS questions were delivered to sections of approximately 150 students. Inquiry-based worksheets were completed in smaller studio sessions of approximately 30-36 students, usually facilitated by a graduate student.

Data Collection

Semi-structured interviews

We conducted semi-structured interviews with the instructors from each course to investigate types of intended student thinking processes elicited by the active learning tasks. We opted to interview course instructors, instead of students, because while there are numerous ways different students might solve the same problem, we are interested in the type of thinking the instructor intended to elicit. We also realized that interviewing students would be methodologically challenging when comparing across different courses and student populations. Our first interviews focused specifically on ARS questions, while subsequent interviews focused on inquiry-based worksheets. The interviews were structured as one hour "think aloud sessions," allowing the interviewees ample opportunity to say whatever came to mind as they were reasoning through a problem or responding to a question. Our interview protocol focused on elements of instructor reasoning and instructional decision making. We began each interview with the following prompt: "I want you to imagine you're looking at these questions from a student's perspective, and I want you to talk through how you would answer each one." We developed this interview prompt as a way to gain insight into how instructors expected students to answer questions. After the instructors worked through the questions from a student perspective we followed up with the prompt: "What was your rationale when assigning this question?" We used this prompt to better understand their instructional reasoning related to each tool. An example of the general interview protocol we used for these interviews is located in Appendix B. We provided instructors with hard copies of the individual ARS questions (or worksheets) to write out their solution processes as they thought aloud. We audio recorded the interviews and collected instructors' written solutions for analysis.

Selection of ARS questions for the biology think aloud interview.

Students in the biology course answered between two and five ARS questions during a 50-minute guided inquiry session, over the course of nine weeks, for a total of 32 ARS questions in the term under investigation. We decided it was feasible, given the brevity of the questions, to have the instructor work through all 32 ARS questions in an hour-long interview.

Selection of ARS questions for the engineering think aloud interview.

Students in the engineering course answered between two to four ARS questions in a 50 minute time period, over the course of 9 weeks, for a total of 33 in the 10-week term under investigation. We decided it was not feasible to expect the instructor to work through all 33 ARS questions in an hour-long interview. As we wanted our sample of questions to be as diverse as possible, we chose 15 of the 33 questions to present to the instructor. Criteria used for choosing questions included the difficulty of the question (which was determined by the percent correct scores from the students), the topic of each question (based on the topics covered in the course outlined in the syllabus), how the percent correct changed if peer instruction was involved, and, if applicable, whether the answer would change if the scenario in the question was altered. Table 5.4, located in Appendix A, details the questions chosen for the interview with the engineering instructor as well as an explanation of why it was chosen, what topic the question covers according to the syllabus, and the percent correct for each question.

Selection of inquiry-based worksheets

For the inquiry-based worksheet interviews we asked the instructors to engage in reasoning through selected worksheets. We selected a single inquiry-based worksheet for the think-aloud interview with the biology instructor. The worksheet focused on parameters of vascular function. We chose this worksheet based on the instructor's input that it was representative of the type of worksheets students would encounter during a guided inquiry session for her course. For the engineering class, we chose the third and ninth worksheets. We chose the third worksheet because it was the first worksheet that used the concept of performing a material balance; this concept is then incorporated into almost all of the subsequent inquiry-based worksheets in the class. We chose the ninth worksheet because it was relatively different from the other worksheets in the class; it required the students to use Excel to perform calculations and a majority of the questions involved qualitative answers. We were curious if the two worksheets would elicit different responses from the instructor. As with the ARS question interviews, we first asked the instructors to solve the worksheets as if they were students encountering the worksheet for the first time and subsequently asked them to explain their rationale for assigning each worksheet.

Data Analyses

We transcribed ARS and inquiry-based worksheet interviews verbatim and analyzed interviewee responses using emergent coding. This process was used to infer categories of meaning and determine the general intended thinking processes that occurred during the interviews. We examined and subsequently broke down each response the instructors gave into the different steps they used to solve a problem. We then generated coding categories based on these process solving steps; we wanted to make these categories as general as possible to better compare between the courses.

Findings

ARS Questions

Our findings indicate the way in which ARS questions engaged students in problem solving or demonstrate understanding varied considerably between the courses. The engineering class presented ARS questions that afforded students the opportunity to apply learned concepts to new scenarios towards improving students' conceptual understanding. In the biology course ARS questions were used mainly to assess if students were correctly interpreting and understanding worksheet questions or to ask students to recall material recently introduced in the worksheet. Another indicator of the different nature of the ARS questions is reflected in the percent of correct responses. A display of the student performance data for the ARS questions is shown in Figure 5.1. In the engineering course, students' first ARS responses averaged 59% correct (labeled ENGR Pre-PI), while in the biology course the average percent correct score was above 89%. The engineering instructor often performed an adaption of Mazur's Peer Instruction (PI) pedagogy[21] and in cases where that was used the second response from the individual students is shown (labeled ENGR Post-PI). There was an average increase of 21% correct responses as a result of PI.

Table 5.1 contains codes that emerged from analysis of the cognitive actions from the ARS interviews performed with the biology and engineering instructors. Tables 5.5 and 5.6, located in Appendix A, contain a description of each code for biology and engineering, respectively. These tables also provide sample interview quotes for each code. Inspection of Table 5.1 shows that the engineering ARS questions were intended to more commonly elicit meaning making from students and also displayed a broader range of required thinking. The biology ARS questions tended to focus more on recall.

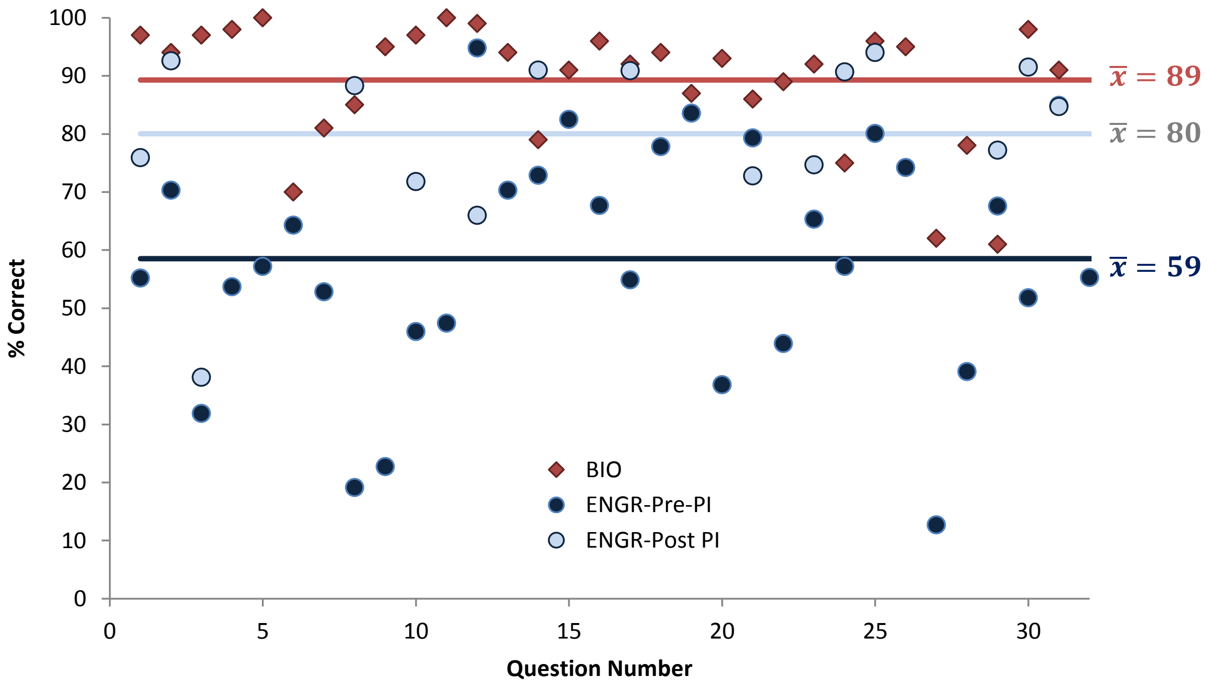


Figure 0.1 ARS question student performance data

Table 0.1 Codes and frequency for ARS questions in engineering and biology courses. Definitions of codes are presented in Appendix A.

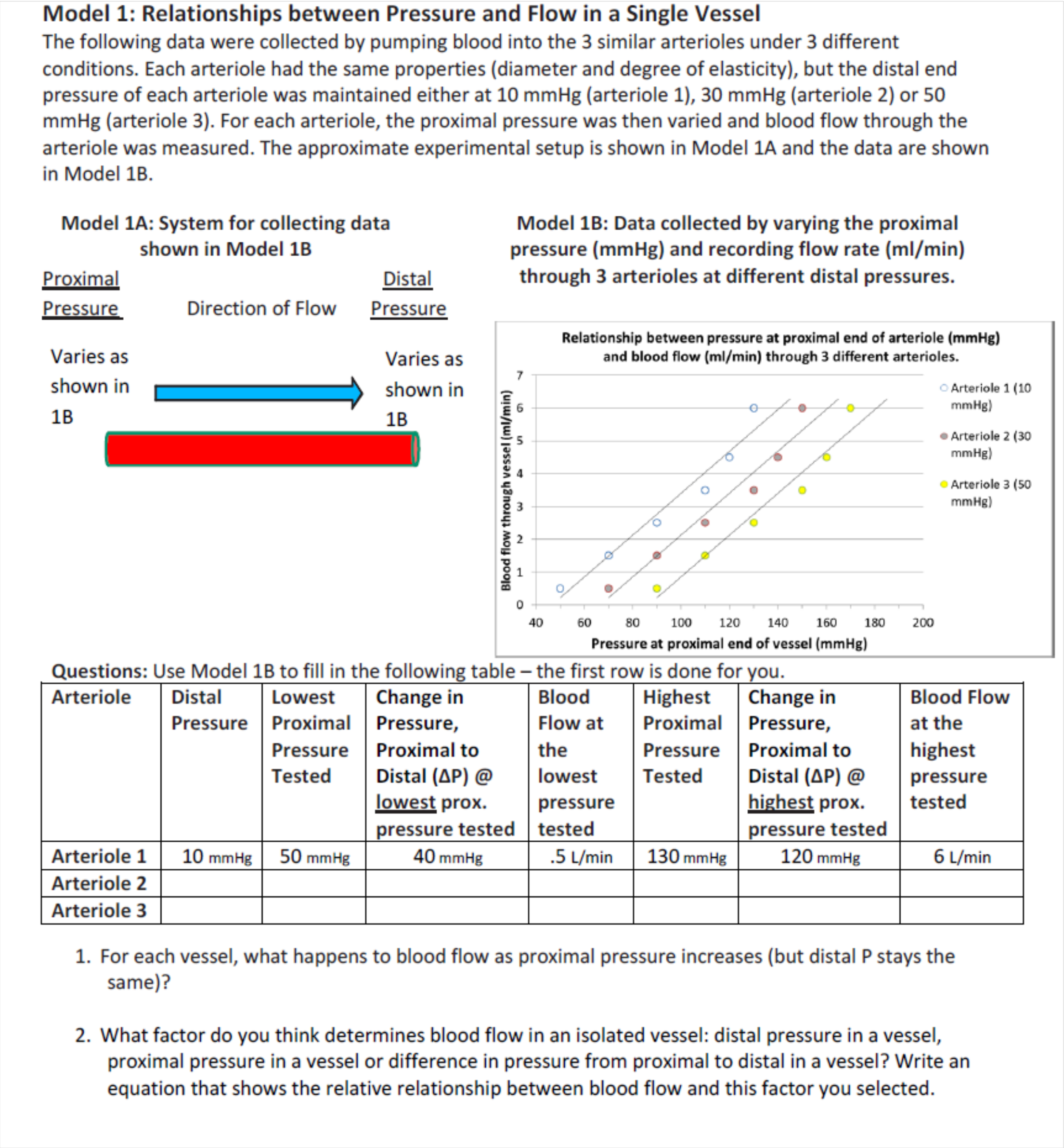
Code	ENGR	BIO
	Frequency (n=15)	Frequency (*n*=32)
Read question completely	11	30
Select information from problem	7	
Immediate recall	1	24
Prior recall	6	3
Recognize concept(s) from problem	11	5
Compare available answers for best choice	3	7
Apply fundamentals of identified concept	2	
Identify relationships between variables	2	
Identify relevant parts of the diagram	3	
Relate identified concept to graphical information	1	
Develop equations to describe phenomena	4	
Rearrange equation to identify relationships between variables	3	
Perform calculation	5	

Inquiry-Based Worksheets

As we found with the ARS questions, instructors utilized inquiry-based worksheets differently. For example, the IBWs were used in the engineering course as a less self-contained structure where students often needed to reference their lecture notes and rely on their knowledge of fundamental principles in order to successfully complete a worksheet. In contrast, in the biology course IBWs were primarily stand-alone activities, containing all of the information necessary for students to answer the questions within the worksheet. Typically the information was presented in a context that aligned with a disciplinary model.

Biology

In the biology course, inquiry-based worksheets engaged students in interpreting models of an underlying concept as well as data sets that students used to answer questions. As described by the instructor the worksheets were intended to guide students through information, enabling them to engage more with the content and make conclusions while developing scientific skills (e.g., data interpretation, critical thinking). The specific inquiry-based worksheet analyzed contained three different models; each model was followed by 2-9 questions students answered based on information from the models. The worksheet ended with six “extension questions” the students answered outside of class time. Although these extension questions required more critical thinking than the previous questions, they still typically referenced the models contained in the worksheet. An example of an inquiry-based worksheet used in the biology course is shown in Figure 5.2.



Model 1: Relationships between Pressure and Flow in a Single Vessel

The following data were collected by pumping blood into the 3 similar arterioles under 3 different conditions. Each arteriole had the same properties (diameter and degree of elasticity), but the distal end pressure of each arteriole was maintained either at 10 mmHg (arteriole 1), 30 mmHg (arteriole 2) or 50 mmHg (arteriole 3). For each arteriole, the proximal pressure was then varied and blood flow through the arteriole was measured. The approximate experimental setup is shown in Model 1A and the data are shown in Model 1B.

Model 1A: System for collecting data shown in Model 1B

Model 1B: Data collected by varying the proximal pressure (mmHg) and recording flow rate (ml/min) through 3 arterioles at different distal pressures.

Questions: Use Model 1B to fill in the following table – the first row is done for you.

Arteriole	Distal Pressure	Lowest Proximal Pressure Tested	Change in Pressure, Proximal to Distal (ΔP) @ lowest prox. pressure tested	Blood Flow at the lowest pressure tested	Highest Proximal Pressure Tested	Change in Pressure, Proximal to Distal (ΔP) @ highest prox. pressure tested	Blood Flow at the highest pressure tested
Arteriole 1	10 mmHg	50 mmHg	40 mmHg	.5 L/min	130 mmHg	120 mmHg	6 L/min
Arteriole 2							
Arteriole 3							

1. For each vessel, what happens to blood flow as proximal pressure increases (but distal P stays the same)?

2. What factor do you think determines blood flow in an isolated vessel: distal pressure in a vessel, proximal pressure in a vessel or difference in pressure from proximal to distal in a vessel? Write an equation that shows the relative relationship between blood flow and this factor you selected.

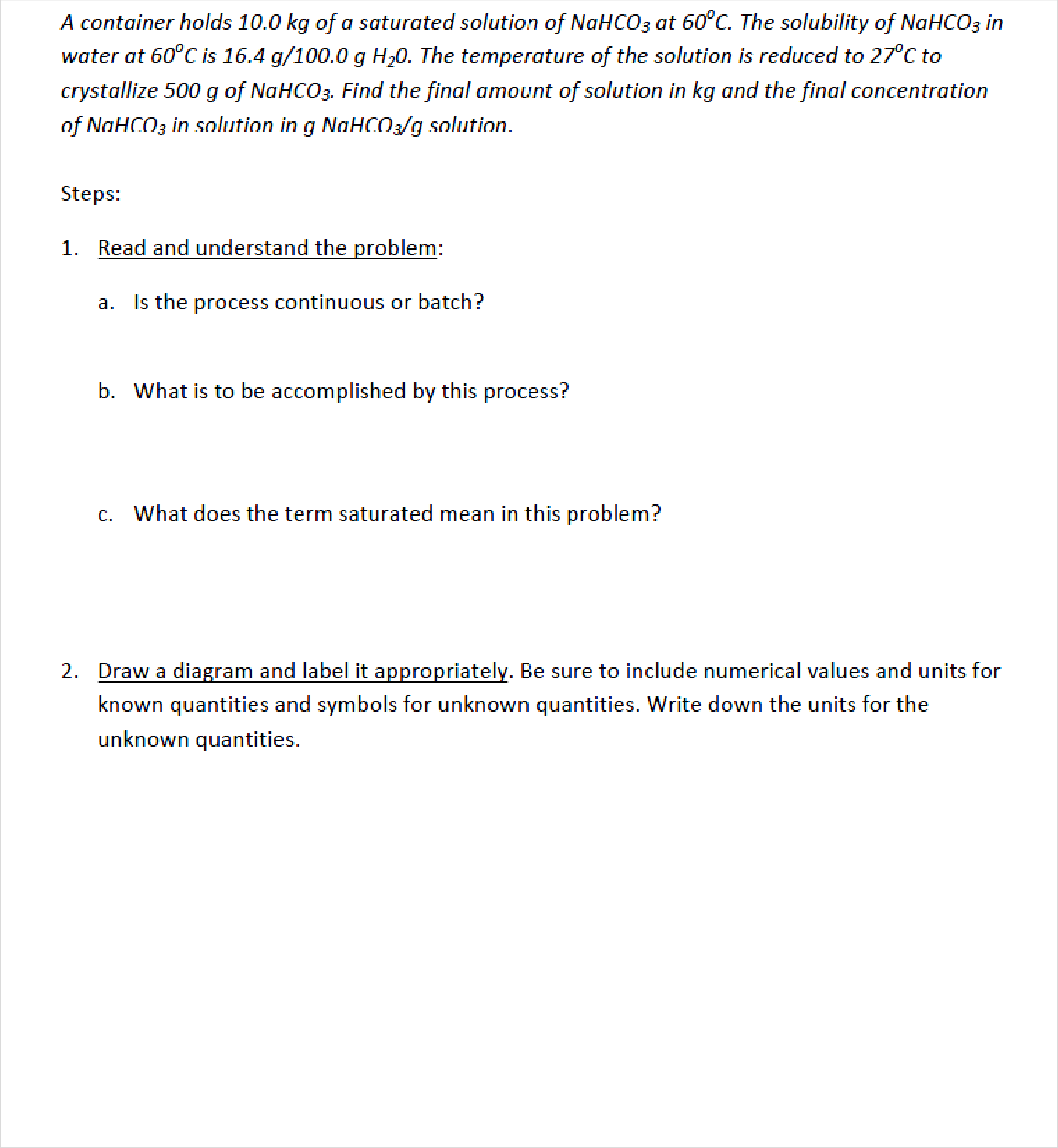
Figure 0.2 Example of biology inquiry-based worksheet question

Engineering

As expressed by the engineering instructor, the inquiry-based worksheets are tied closely to information introduced in lecture. Most worksheets involve a main problem that includes a problem statement with relevant information students need to use to solve the worksheet. The problem is then broken down into different steps that students complete to help them reach an answer. These steps are modeled after problem solving processes introduced in lecture although the inquiry-based worksheets are less scaffolded than examples students work through in lecture.

The inquiry-based worksheet selected from week 3 starts with a problem statement followed by several questions which require students to read and interpret the problem statement and a question that asks then to draw and label a diagram that represents what is happening in the problem. The next section walks students through a scaffolded problem solving process similar to what they covered in previous lectures. The final section of the worksheet contains two reflection questions that ask the students to relate the original problem statement to real world engineering scenarios.

The inquiry-based worksheet selected from week 9 begins with a qualitative reasoning section that contains several conceptual questions related to topics recently discussed in lecture. This is succeeded by a problem solving section which requires students to write and subsequently rearrange equations, utilize the Solver function in Excel, and then answer questions (both qualitative and quantitative) based on the values they obtain from Excel. An example of an inquiry-based worksheet used in the engineering course is shown in Figure 5.3.



A container holds 10.0 kg of a saturated solution of $NaHCO_3$ at 60^{o}C. The solubility of $NaHCO_3$ in water at 60^{o}C is 16.4 g/100.0 g $H_2$0. The temperature of the solution is reduced to 27^{o}C to crystallize 500 g of $NaHCO_3$. Find the final amount of solution in kg and the final concentration of $NaHCO_3$ in solution in g $NaHCO_3$/g solution.

Steps:

1. Read and understand the problem:
    a. Is the process continuous or batch?
    b. What is to be accomplished by this process?
    c. What does the term saturated mean in this problem?
2. Draw a diagram and label it appropriately. Be sure to include numerical values and units for known quantities and symbols for unknown quantities. Write down the units for the unknown quantities.

Figure 0.3 Example of engineering inquiry-based worksheet questions

Table 5.2 contains the coding results for the inquiry-based worksheets. The extension questions on the biology worksheet were not coded because they were designed to be done outside of class. Inspection of Table 5.2 shows that the engineering worksheets require students to reference previous knowledge and information presented in lecture more than the biology worksheet. The biology worksheet is more self-contained in the sense that most of the information students need to complete the worksheet is provided via the models.

Table 0.2 Count of emergent themes in engineering inquiry-based worksheet interview

Code	Biology	Engineering	
	n=19	Week 3 n=8	Week 9 n=12
Recall given info	18	11	2
Recall previous answer	9	4	5
Recall previous knowledge	4	12	5
Recall lecture info	2	7	5
Relate info to physical representation		1	
Transfer graphical info into numerical information	4		
Use graphical info to qualitatively explain situation	8		
Rearrange equations		1	4
Create equation	4		
Explain answer/critical thinking	6		
Reflect		1	1
Use Excel			7

Presentation of Cases

A summary of the differences in the context and use of the active teaching tools between engineering and biology is presented in Table 5.3. Below we provide a more detailed description of each class including how the active teaching tools fit together in relation to other important elements of the courses.

Table 0.3 Alignment across different elements of the two courses

	Biology			Engineering		
	Frequency	Size	Purpose	Frequency	Size	Purpose
Lecture	Twice a week, 50 minutes	165 students	Content is introduced and students engage in ARS questions and discussions	Twice a week, 50 minutes	300 students	Content is introduced and transmitted to students
ARS Questions	Three times a week	165 students	"Check in" with students to see if they are correctly interpreting worksheet content or ask students to report their understanding of worksheet questions just completed.	Once a week, 50 minutes	Approx. 150 students/ section	Strengthen conceptual understanding by building on topics from lecture
Inquiry-Based Worksheet	Once a week, 50 minutes	165 students	Guide students through biological models, enabling them to engage with the content using disciplinary thinking	Once a week, 50 minutes	Approx. 30 students/ section	Scaffolded worksheets reinforce problem solving skills/processes instituted in lecture

Biology

The biology course, *Advanced Anatomy and Physiology*, is the final course in an upper division 3-term sequence series required for biology majors. Prerequisites for the course included the completion of the preceding courses in the year-long sequence and concurrent or previous enrollment in the affiliated lab section. Based on the instructor's records, this course had an enrollment of 165 students during the term in which our data was collected. Classes met on Monday, Wednesday and Friday for 50 minutes. Monday and Wednesday classes were instructor-led, lecture-based and attended by all students. Friday of each week students engaged in guided inquiry sessions that were instructor-led, but facilitated by seven trained undergraduate learning assistants.

During lecture periods (M, W) and guided inquiry sessions (F), ARS questions were presented to students as PowerPoint slides which students answered for extra credit points using in-class response clickers (Turning Technologies). We focused on ARS questions from the guided inquiry sessions to be able to compare to the studio sessions in the engineering course. Students answered ARS questions collaboratively and individually and were generally allowed 1-minute to respond.

Collective responses were displayed to the class and discussed, depending on the percent correct, prior to moving on. The number of questions asked during each session varied between two and five. A total of 32 questions were asked throughout the term.

The biology course implemented inquiry-based worksheets during the guided inquiry sessions on Fridays. A typical Friday session (50 minutes) consisted of an introduction (2 minutes), student group work facilitated by undergraduate learning assistants (15-25 minutes), ARS questions and group discussion (5-10 minutes), continued student group work (10-15 minutes) and wrap up (2-5 minutes). During these sessions the classroom was divided into seven territories with a learning assistant assigned to each territory. The instructor assigned students to groups of three which were maintained for the duration of the term. Students collaboratively worked on inquiry-based worksheets answering an average of 19 worksheet questions during a guided inquiry session. Worksheets also contained "extension questions" that students engaged in outside of class. Extension questions consisted of three types, classified by the instructor as (i) *big picture questions*, consisting of open-ended questions to stimulate discussion and look at larger concepts, (ii) *detail questions* about course content and (iii) *clinical correlation questions* situating concepts in clinical applications.

Engineering

The engineering course is the first department-specific prerequisite in the curriculum and is followed in series by *Energy Balances* and *Process Data Analysis*. Based on the number of students who took the final exam, this class had an enrollment of 287 students during the term in which our data was collected. This course began and ended each week with two 50 minute, instructor-led lecture sections held on Monday and Friday. All 287 students attended the same lecture. Following the Monday lecture section, students participated in one of two weekly, 50 minute recitations held on Tuesday or Wednesday. Each recitation section, which was also taught by the instructor, contained approximately half of the total enrolled students. Following the recitations, the students also attended a 50 minute studio session once a week on either Wednesday or Thursday. Each studio contained approximately 30-36 students. Out of ten total studio sessions, one honors session was run by the instructor and the rest were facilitated by graduate students.

Within the engineering course, the ARS questions were presented to students during the recitation sections via the *AIChE Concept Warehouse*.[22] Students responded to the questions using their laptops, smartphones, or tablets and the answers were stored in a database. Along with their multiple choice answer, students were asked to provide an explanation of their answer as well as a confidence score between one and five to determine how sure they were of their answer. The instructor commonly implemented a peer instruction pedagogy into the recitation sessions. For non-peer instruction questions, the instructor assigned a question, students answered it individually, and then the instructor displayed the responses and discussed the question. When peer instruction was implemented, the instructor re-assigned the same question after the students answered it individually and students discussed the questions with their neighbors. After the students conferred with one another they independently answered the question again. The number of questions asked during each session varied between two and four. A total of 33 questions were asked throughout the term.

The engineering course implemented the inquiry-based worksheets during the studio sessions. Some worksheets involved an introduction section that was to be completed individually, while others solely contained group work components. Students were placed into small groups by the session facilitator (either the instructor or, more likely, a graduate student). Studio groups typically contained three students, although, depending on the amount of students in the studio, some groups contained four students. Groups were kept consistent for 5 weeks, after which students were assigned to new groups by the facilitator for the remainder of the term. Graduate students were instructed to respond to student questions with subsequent guiding questions, as opposed to providing students with answers to the worksheets. There were a total of nine inquiry-based worksheets completed throughout the term. Students were given 50 minutes to work on the worksheets which were collected, regardless of completeness, at the end of each studio session.

Discussion

As evidenced in our findings, variations in the ways active teaching tools are designed and enacted produce different types of thinking processes and engagement with course content. Although we cannot speak for the instructors as to why they chose to implement the active learning tools in different ways, during the think aloud interviews we asked the instructors questions regarding their rationale for using the ARS questions and inquiry-based worksheets in their courses. We hoped their answers would help illuminate why they implemented the active learning tools in the ways that they did. For example, during the inquiry-based worksheet interview the engineering instructor described how he uses both tools in relation to how he approaches lecture. This quote helped us understand his intention behind how he implemented the two tools:

> **Engineering instructor:** ... trying to have a cohesive, kind of weekly routine of content delivery: reinforced conceptual understanding in recitation, scaffolded application in studio. So kind of the idea those two lectures, the recitation, and the studio as being like a learning unit...and then the homework follows that... my lecture has become more content delivery…"

This overview of how the instructor intends for the tools to fit together in the engineering course, as well as several other quotes from the instructors regarding their rationale for structuring these tools, gave us insight into why the implementation differences we discovered in this study existed. In this section we will highlight and discuss other instructor quotes to provide potential explanations as to why these differences occurred.

ARS Questions

When questioned about her rationale behind designing ARS questions, the biology instructor acknowledged that they can be a helpful tool to use in large classes.

> **Biology Instructor:** ...and so you know the value of clicker questions in rooms of greater than 30 people is…: In a really big class I can't see their sheets, and so I don't know what they're thinking. And it's really useful to check in with them in that way.

She pointed out that she often uses the ARS questions to check in with students and ensure they're following along.

When asked about how he wants students to be engaged while solving ARS questions the engineering instructor explained he wanted to push students beyond procedural problem solving.

> **Engineering Instructor:** I guess trying to get them to start to create that knowledge structure in their head. That there are certain conventions and there are certain cues that are gonna help

> them bend those problems into, you know, help them find a solution...Trying to provide cues that are similar to things they're gonna see in other parts of the class, homework and exams and so forth, to get them to hone in on those specific concepts and then in some cases manipulate or examine at a level that they're not gonna get just by plugging and chugging into those equations.

He expressed how he wants students to do more than solve a simple calculation; instead, he wants them to work with and be able to manipulate equations as well as focus on the concepts involved in each question.

Inquiry-Based Worksheets

When asked why she recommend we use a specific inquiry-based worksheet for the think aloud interview, the biology instructor explained that she thought it was the epitome of a typical inquiry-based worksheet for the class.

> **Biology Instructor:** And so what I really like about POGIL that this worksheet adheres to is you can get everything you need from this, you know, strictly from the models, and your brain, and thinking about things. And maybe if you don't know what these vessels are, yeah, you could look them up, but you probably do, you know, based on where my students are at. And so, like that's what I like. This is very much a standalone.

Here she expresses how the inquiry-based worksheets in the biology course are self-contained; students can complete the worksheets with little or no prior knowledge. This is a direct contrast with the structure of the inquiry-based worksheets in the engineering class.

When prompted about why he uses studio sessions in his class, the engineering instructor described the relationship between the inquiry-based worksheets and other aspects of the course.

> **Engineering Instructor:** I view studio as a really scaffolded and supported place for students to have their first experience applying the principles from lecture. So you kind of get all this information and not a lot of chances to engage it in lecture and before you get a blank problem statement from the homework assignment and are left with a blank page [you get a chance] to walk through the steps or the concepts that are gonna have to be applied as we move forward to homework and exams. Having it be a place where they've got classmates they can bounce ideas [off of]…"

Here he clarifies that he views the inquiry-based worksheets as a step students can use in between learning ways to solve problems in lecture and applying these problem solving methods on homework assignments and exams.

Both instructors adapted and enacted the ARS questions and inquiry-based worksheets to fit the structure of their courses. The engineering instructor used the ARS questions to reinforce conceptual understanding and he wanted the questions to engage the students beyond basic procedural problem solving. The biology instructor used the ARS questions as a way to check in with her students and assess whether or not they were following along. The engineering instructor used the inquiry-based worksheet to build problem solving skills and saw them as a less self-contained way for students to work with information from lecture. In fact, the intent was to have students make connections to lecture. The biology instructor used the inquiry-based worksheets to guide students and allow them to engage more with the content and make conclusions.

Conclusion

On the surface these findings may seem apparent; two different instructors used similar tools in different ways according to their class goals. Although both instructors that participated in this study articulated the goals of their courses, it is unclear if they thought of their goals in respect to other choices. Did they consider alternative tools or implementations when they planned the structure of their courses? Any instructor looking to implement active teaching tools should be aware of the flexibility of the tools they intend to use as well as the goals they hope to accomplish. Not all active teaching tools are the same. We know studies show they are more effective than passive lectures alone, but other than the overall effectiveness of these tools there has been little investigation into their variability and how they fit into different classrooms.

This study also serves as an in-depth look at how two tools were implemented in different ways which future instructors could reference when deciding how they plan to implement similar active learning tools. In the future we want to analyze other courses that have implemented ARS questions and inquiry based worksheets for an even richer comparison. We are working towards the ultimate goal of creating a university wide “tool map” that details which tools are implemented in each department as well as detailed information about how those tools fit into each course. This would serve as an inventory of both the student-centered learning activities currently being used on campus as well as how the implementation of these tools vary across instructors and courses.

References

1. Arum, R., & Roksa, J. (2011). Academically adrift: Limited learning on college campuses. University of Chicago Press.
2. National Research Council, Discipline-based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering, S.R. Singer, N.R. Nielsen, and H.A. Schweingruber, Editors. 2012: Washington, DC
3. Maltese, A. V., & Tai, R. H. (2011). Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among US students. Science Education, 95(5), 877-907.
4. Museus, S. D., & Liverman, D. (2010). High performing institutions and their implications for studying underrepresented minority students in STEM. New Directions for Institutional Research, 2010(148), 17-27.
5. Shapiro, C. A., & Sax, L. J. (2011). Major selection and persistence for women in STEM. New Directions for Institutional Research, 2011(152), 5-18.
6. Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. American Journal of Physics, 66, 64-74. doi: 10.1119/1.18809
7. Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. Science, 332(6034), 1213-1216.
8. President’s Council of Advisors on Science and Technology (PCAST). (2012). Engage to Excel: Producing One Million Additional College Graduates With Degrees in Science, Technology, Engineering, and Mathematics. Washington, DC: Author.
9. Klymkowsky, M. W., Garvin-Doxas, K., & Zeilik, M. (2003). Bioliteracy and teaching efficacy: what biologists can learn from physicists. Cell Biology Education, 2(3), 155-161.
10. Chi, M. T. (2009). Active constructive interactive: A conceptual framework for differentiating learning activities. Topics in Cognitive Science, 1(1), 73-105.
11. Fairweather, J. (2008). Linking evidence and promising practices in science, technology, engineering, and mathematics (STEM) undergraduate education. Board of Science Education, National Research Council, The National Academies, Washington, DC.
12. Bonwell, C. C., & Eison, J. A. (1991). Active Learning: Creating Excitement in the Classroom. 1991 ASHE-ERIC Higher Education Reports. ERIC Clearinghouse on Higher Education, The George Washington University, One Dupont Circle, Suite 630, Washington, DC 20036-1183.
13. Prince, M. (2004). Does active learning work? A review of the research. Journal of Engineering Education, 93(3), 223-231. doi: 10.1002/j.2168-9830.2004.tb00809.x
14. McManus, D. A. (2001). The two paradigms of education and the peer review of teaching. Journal of Geoscience Education, 49(5), 423-434.
15. Chi, M., T., H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. Educational Psychologist, 49:4, 219-243, doi: 10.1080/00461520.2014.965823
16. Wood, W.B. (2009). Innovations in teaching undergraduate biology and why we need them. Annu. Rev. Cell Dev. Biol. 25, 93-112

17. Smith, M. K., Wood, W. B., Krauter, K., & Knight, J. K. (2010). Combining peer discussion with instructor explanation increases student learning from in-class concept questions. CBE Life Sciences, 10, 55-63.
18. Armbruster, P., Patel, M., Johnson, E., & Weirss, M. (2009). Active learning and student-centered pedagogy improve student attitudes and performance in introductory biology. CBE Life Sciences Education, 8, 203-213.
19. Ebert-May, D., Derting, T. L., Hodder, J., Momsen, J. L., Long, T. M., & Jardeleza, S. E. (2011). What we say is not what we do: effective evaluation of faculty professional development programs. BioScience, 61(7), 550-558.
20. Borrego, M., Cutler, S., Prince, M., Henderson, C. & Froyd, J. E. (2013). Fidelity of implementation of research-based instructional strategies (RBIS) in engineering science courses. Journal of Engineering Education, 102, 394–425. doi: 10.1002/jee.20020
21. Mazur, E. (1997). Peer instruction (pp. 9-18). Upper Saddle River, NJ: Prentice Hall.
22. Koretsky, M., Falconer, J., Brooks, B., Gilbuena, D., Silverstein, D., Smith, C., & Miletic, M. (2014). The AIChE Concept Warehouse: A tool to promote conceptual learning. Advances in Engineering Education, 4, 89.

Appendix A– Questions and Coding Data from ARS Questions

Figure 0.4 Explanation of questions selected for ARS guided inquiry session for engineering course

Question Title	Explanation	Topic from Syllabus	%Correct 1	%Correct 2
Ice water displacement	One of the more in-depth questions/requires a lot of reasoning. Also an example of a question where the answer changes if the situation is adjusted (ex: if the ice is distributed throughout the water)	Process and Process Variables	23.7	34.7
Compositions from separator	Genuinely curious how the instructor expected students to solve this question	Law of Conservation of Mass	47.9	88.9
Formaldehyde Production Problem II	Used to ensure students are understanding concepts (fractional conversion, yield, etc.). Also used as interpretation questions. Instructor thought process for these questions could be applied to other "easy questions." Chose this one because of the % correct. Relatively high even on first attempt	Material Balance on Reactive Systems	69.6	88.7
Mass Flow Rate Through Conic Section II	These questions were asked back to back. Interested in instructor thought process/perceived student thought processes about having them in sequence	Process and Process Variables	49.3	72.2
Mass Flow Rate Through Conic Section IV			70.3	96.7
Ideal Gas Law, Effect of Temperature on Isobar	A lot of the Single Phase questions are a matter of using the ideal gas law and manipulating it to derive an answer. Asking about one of them could give insight into the other ones as well. Chose this one because of the high percent correct	Single Phase	86.2	---
Two-Unit System DOF II (Unit 2 Balance)	Several different DOF questions. This has more content than students simply getting confused by a splitting point. Can hopefully be applied to other DOF questions	Balances on Multiple-Unit Processes	35.1	50.4

Relative humidity and the time of the day	Creative question and curious about instructor reasoning behind creating it	Multiphase Systems	50	87.2
Continuous Reactor Molar Balance	This is a very easy question. Curious if the instructor presented it as a check	Material Balance on Reactive Systems	78.9	---
Pressure Change After Gas-Phase Reaction	Chosen for high delta of percent correct	Single Phase Systems	36.6	79.5
Splitting Point Molar Compositions	Seems like an easy question/concept yet most students got it wrong. Chosen for juxtaposition between perceived easiness and the low percent correct	Material Balances with Recycle and Bypass Streams	12	---
Calculating Equilibrium Gas Pressure Between Two Tanks	Relatively high delta and another ideal gas law question. Could be compared to others that involve ideal gas law.	Single Phase Systems	51.4	90.9
General Energy Balance for Molar Species in Reaction	Based on percent correct this was an "easy question". They had similar questions on exams, specifically extra credit sections. Curious how the instructor wanted them to think through questions like this. It seems like a matter of remembering definitions more than solving a problem	Material Balance on Reactive Systems	94	---
Linear Independence	Students seemed to have a hard time grasping linear independence. Curious about what the instructor has to say about this question, it is the only one that directly addresses this concept	Law of Conservation of Mass	50.9	---
Two-Phase Gibbs Phase Rule I	Chosen for difficulty and as another "multiphase" question	Multiphase Systems	11	---

Figure 0.5 Types of thinking processes elicited from biology ARS questions used during guided inquiry sessions

Code	Description	Example of Thinking Process	Frequency (*n*=32)
Read question completely	The entire question and related content was read prior to working through solution	Not applicable	30
Immediate recall	The question could be answered from work completed during the guided inquiry session	"So to answer this, I think I would look to my worksheet, so whatever that model is. So that model with the, oh it's graphs as I recall, so model with graphs, and then I would reference my conversation that I had."	24
Prior recall	Answering the question involved recalling information from a prior class	"I think I would answer it by defining what I know about the P wave, the QRS and the T waves, and do I know where they are? Do I know that they're in the EKG? I think that I would know that the P, QRS and T wave belong to the EKG. And then I would probably draw for myself the EKG."	3
Compare available answers for best choice	Available multiple choice answers were reviewed to decide which answer made the most sense	"I think that I would have to say okay, you know, we just went through these four questions from this worksheet, and, you know, we eliminated some of these options basically as being the right answer."	7
Recognize concept(s) from problem	The question could be answered by recognizing concepts and applying worksheet solutions in a slightly different context	"And so I would be thinking to myself okay, I just learned from the one model about pressure and valves in the left side of the heart, and now I'm being asked about the right side of the heart"	5

Figure 0.6 Types of thinking processes elicited from engineering ARS questions used during guided inquiry sessions

Code	Description	Example of Thinking Process	Frequency (n=15)
Read question completely	This was used when the instructor read the entire question before beginning to solve the problem	"Okay, the following volume temperature plots are made at constant pressures, constant numbers of moles of gas of an ideal gas, which plot represents the higher pressure?"	11
Select information from problem	This was used if the instructor chose information from the problem prompt to help figure out how to solve part of the problem.	"…we have composition information on one input stream, two output streams, and there's an empty third output stream, want to know what the composition is going to be…	7
Recognize concept(s) from problem	This code applied when the instructor used the main concept involved in a problem to determine what needed to be done to solve the problem	"Okay, I'm thinking, again, degrees of freedom, but we've got multiple phases here and so we need to use Gibbs phase rule…"	11
Develop equations to describe phenomena	This code was used when the instructor developed, as opposed to simply remembered, equations to describe what was happening in the problem	"And in this case I want to do a material balance on C… I'm just gonna write it out, in minus out plus generation minus consumption equals zero…"	4
Compare available answers for best choice	This code described when the instructor looked at the available multiple choice answers and compared them to decide which answer made the most sense	"So I 'm gonna look for one that has generation, consumption and no accumulation, and that one, I see, is this fourth one down…"	3
Apply fundamentals of identified concept	After the instructor identified a concept, he often used the fundamentals of that concept to help solve the problem. This code represented that	"…think about the mass flow rate and use that to say that the mass in and out is gonna be the same. So I'm going to, yeah, focus on that concept of mass conservation here."	2
Immediate Recall	This code refers to the instructor referencing a previous problem to help solve the current problem	"…we would have gone through this [first question] and said mass is conserved…"	1
Identify relationships between variables	This code often occurred in conjunction with developed or remembered relationships. The instructor often wrote an equation and then needed to determine how the variables in the equation related to one another	"…if I were a good student I would start with a force balance on just one ice cube sitting in the cup…And then basically breaking down what are the components of that force into mg for the gravitational force and for the buoyant force the displaced volume times the density of water. I think that the link that they have to make here is thinking about once this ice cube melts that it is water."	2
Prior Recall	This code was used when a problem involved a concept with a related equation that the instructor needed to remember and write out or remember know the definition of words in the problem statement in order to solve the problem	"So, I'm given information about volumes, temperatures, and pressures; that makes me think ideal gas law. So I'm gonna write that out."	6

Rearrange equation to identify relationships between variables	This code was used to describe whenever the instructor would rearrange an equation he was using	"…so if I write the ideal gas law, PV=nRT, then the number of moles of gas in that system, and that's gonna be PV over RT and I'll call that n_1 and we'll call that V_1 and in the other one we've got n_2, and we'll call it P_1 and P_2 too, V_2 over RT...."	3
Identify relevant parts of the diagram	This code was used when the instructor related information from the problem prompt to parts of a diagram	"So when it says recovered and recycled I'm gonna look for somewhere where methanol is gonna be separated and sent back to the beginning of the process. So I can see that's the unit that's labeled S here."	3
Relate identified concept to graphical information	This code was used when the instructor used a graphical representation to think through a concept. The problem did not specifically require a graph to be drawn but the instructor did so while working through the problem	"So I'm gonna think about a phase diagram for water and think about the fact that the relative humidity is going to tell me something about the fraction of the saturation temperature that I'm at."	1
Perform calculation	This code was used whenever a numerical calculation was required	"…And for C my in is 50, my out is what I'm looking for, my generation is plus 2 times the extent of reaction and there's no consumption. Okay, and so my output is gonna be 50 plus 2 times the extent of reaction."	5

Appendix B- Interview Protocol

Research Rationale:

To better understand the rationale and intention of instructor presented [ARS/inquiry-based worksheet] questions in student-centered active learning environments in STEM disciplines.

Data collection goals:

Understand a faculty member's use of questions in active learning environments.

Intro script:

Is it OK if I record this conversation?

[State date and name of interviewee]

[Guidelines for what to say, does not need to be exact]

Thank you for taking the time to talk with me today. The focus of this interview is on "student thinking processes" as students complete [ARS questions/inquiry based worksheets]. I’ve chosen some [questions/worksheets] you’ve previously used in class that I want to discuss with you today.

During this interview, I want you to imagine you’re looking at these questions from a student’s perspective, and I want you to talk through how you would answer each one. I have the questions printed out; feel free to make notes as you’re thinking aloud.

Prompt 1: Imagine you are a good student who has just been assigned this [question/worksheet], talk through how you would answer it

Probe: That’s your design, how would a student think of it?
Probe: Remember, you’re looking at this question as a student.
Probe: Why did you say that? Why did you make that choice?
Probe: What is your final answer?
Probe: How would you solve the question from here?

[Repeat for all questions]
Prompt 2: What was your rationale when assigning this [question/worksheet]?
Probe: Why did you choose this worksheet/how does it fit into what or how you wanted students to learn?
Probe: What other activities does it build on?
Probe: Can you think of anything specific about this question that might confuse students?
[Repeat for all questions]

After prompts:

What is your rationale for using [ARS questions/inquiry-based worksheets] in your class?
When you create a [question/worksheet] how do you structure it? What’s the creative process like?

What do you want students to gain from a typical [ARS question/inquiry-based worksheet]?

What is your approach as a facilitator in these active learning environments?

6. GENERAL CONCLUSION

The studies depicted in the preceding four manuscripts investigate the different roles active learning tools play in engineering classroom learning environments while providing an in-depth look at the thinking processes each tool elicits from students. While the first three studies focus on important elements within active learning, the fourth study focuses more on how those elements fit together within a course.

In the first two metacognitive studies we asked students to reflect on their own learning; we then examined how using different reflection prompts affected the responses students submitted. The third study is a cognitive examination of how students think through virtual laboratories as well as an attempt at understanding how common misconceptions manifest themselves in student thinking processes. The fourth study focuses on context and how different active learning tools look in different course settings.

Bibliography

1. Prince, M. (2004). Does active learning work? A review of the research. Journal of Engineering Education, 93(3), 23-231. doi: 10.1002/j.2168-9830.2004.tb00809.x
2. Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. Educational Psychologist, 49:4, 219-243, doi: 10.1080/00461520.2014.965823
3. Chi, M. T. (2009). Active constructive interactive: A conceptual framework for differentiating learning activities. Topics in Cognitive Science, 1(1), 73-105.
4. Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66,* 64-74. doi: 10.1119/1.18809
5. Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. PNAS: Proceedings of

the National Academy of the United States of America, 111(23), 8410-8415. doi:10.1073/pnas.1319030111