

James Madison University

JMU Scholarly Commons

Senior Honors Projects, 2020-current

Honors College

5-12-2023

Student-led success: Improving attendance and learning at Fourth Hour review sessions

Andrew C. Gilbert

James Madison University

Follow this and additional works at: <https://commons.lib.jmu.edu/honors202029>



Part of the [Education Commons](#), and the [Engineering Commons](#)

Recommended Citation

Gilbert, Andrew C., "Student-led success: Improving attendance and learning at Fourth Hour review sessions" (2023). *Senior Honors Projects, 2020-current*. 164.

<https://commons.lib.jmu.edu/honors202029/164>

This Thesis is brought to you for free and open access by the Honors College at JMU Scholarly Commons. It has been accepted for inclusion in Senior Honors Projects, 2020-current by an authorized administrator of JMU Scholarly Commons. For more information, please contact dc_admin@jmu.edu.

Student-Led Success: Improving Attendance and Learning at Fourth Hour Review Sessions

An Honors College Project Presented to
the Faculty of the Undergraduate
College of Integrated Science and Engineering
James Madison University

by Andrew Carter Gilbert

May 2023

Accepted by the faculty of the Computer Science Department, James Madison University, in partial fulfillment of the requirements for the Honors College.

FACULTY COMMITTEE:

HONORS COLLEGE APPROVAL:

Project Advisor: Dee A. B. Weikle, Ph.D.
Associate Professor, Computer Science

Bethany Blackstone, Ph.D.
Dean, Honors College

Reader: Chris Mayfield, Ph.D.
Associate Professor, Computer Science

Reader: Chris Johnson, Ph.D.
Associate Professor, Computer Science

Reader: Michael Stewart, Ph.D
Assistant Professor, Computer Science

PUBLIC PRESENTATION

This work is accepted for presentation, in part or in full, at James Madison University on April 21, 2023.

Copyright © 2023 by Andrew C. Gilbert

All Rights Reserved

Dedication

I would like to dedicate this project to all of my friends and family for their support throughout this process. As well as to my advisors Dr. Weikle and Dr. Stewart for all of their guidance and helping motivate me through my first thesis.

Table of Contents

	Page
Table of Contents.....	4
List of Tables.....	5
List of Figures.....	5
Chapter 1: Introduction.....	7
1.1 Purpose and Objectives.....	7
1.2 Background and Literature Review.....	8
1.2.1 Peer Instruction.....	8
1.2.2 Teaching Assistant Led Review Sessions.....	9
Chapter 2: Methodology.....	11
2.1 Student Attendance.....	11
2.2 Content Retention.....	11
2.2.1 Measuring Learning Gains.....	11
2.2.2 Data Analysis.....	12
Chapter 3: Implementation.....	13
3.1 Study Session Design.....	13
3.1.1 Previous Approach.....	13
3.1.2 Current Approach.....	13
3.1.3 Session Schedule and Misconceptions Covered.....	14
Chapter 4: Results.....	16
4.1 Attendance.....	16
4.2 Learning Gains During The Fourth Hour.....	17
4.3 Learning Gains Throughout the Semester.....	21
Chapter 5: Conclusions and Future Work.....	24
5.1 Primary Results.....	24
5.2 Future Work.....	24
Bibliography.....	26
Appendix.....	28

List of Tables

Table	Page
3.1.3 Fourth Hour Schedule and Misconceptions Covered.....	15

List of Figures

Figure	Page
4.1 Fourth Hour Attendance Spring vs Fall 2022.....	17
4.2.1 Week 3 PI Progression.....	19
4.2.2 Week 12 PI Progression.....	19
4.3.1 Exam Grades of Students Who Did vs Did Not Attend Fourth Hour.....	23

Abstract

This research continues the work of Gilbert et al. [1] by translating the “Fourth Hour” review session materials from Java to Python, keeping the focus on common misconceptions in the literature, continuing to use Peer Instruction (PI), and addressing two issues in the previous study: low attendance and short-term learning outcomes during the session. [2] The “Fourth Hour” is a review session designed to support students taking CS1 who may have missed class or need to review the previous week’s material. This review session is an hour long and is offered twice a week with the same material so more students can fit it into their schedules. In Gilbert et al’s study, attendance was too low for the results to be statistically significant. This new study attempts to address this issue by implementing in-class weekly assessments with retakes that will be used to improve attendance. This study also measures learning gains during the session by measuring the accuracy of the responses to the given PI questions to determine whether students who attend are actually learning the material being presented.

Chapter 1: Introduction

As computing becomes more commonplace in our world, society needs more people to apply computing expertise to more disciplines. According to the U.S. Bureau of Labor Statistics, “Employment of computer and information research scientists is projected to grow 22 percent from 2020 to 2030, much faster than the average for all occupations.” [3] This need results in more jobs in computing and more students wanting computing majors. Computer science departments are responding by trying to educate not only more students, but to attract and retain students from underrepresented populations in computing. Many students struggle to grasp fundamental concepts and feel they lack the ability to succeed. At James Madison University, undergraduate learning is our highest priority and our Computer Science program is no different. To ensure students with little to no experience can learn the foundational Computer Science concepts, we have switched our CS1 course from Java to Python to more effectively scaffold students without prior experience into the field and cast a wider net of future computer scientists.

1.1 Purpose and Objectives

The overall purpose of this research is to study and improve the mechanism for students who find themselves falling behind in CS1 or don't have significant prior experience to stay caught up and succeed in the course. The Fourth Hour review session was designed several years ago to address this issue. Because the CS curriculum builds from one week to the next, getting behind can cause serious problems. A weekly review session can help students catch up before the difficulty is insurmountable. We have chosen to revisit the review sessions from Gilbert et al. and recreate them in Python since CS1 is now in Python. We also plan to resolve an issue with this research by improving attendance. This study focuses on revising the materials and measuring

learning gains for each session as well as throughout the semester. Another coordinated study is implementing in-class weekly assessments and retakes to incentivize attendance. As part of this research, we will be investigating the use of review sessions and how attendance impacts exam scores.

The objectives of this study are to update the materials including slides and peer instruction questions for the Fourth Hour to reflect the language change of JMU's CS1 course from Python to Java, repeat the semester-long learning gain analysis in Gilbert et al, and analyze learning and evaluate peer instruction questions for individual review sessions. [1] These research objectives all depend on getting strong attendance and participation during the review sessions. Attendance was incentivized by offering retakes of review quizzes given in individual sections of the class. The effectiveness of this incentive is discussed in a separate study. [8]

1.2 Background and Literature Review

1.2.1 Peer Instruction

Peer Instruction was selected to be used during the review sessions. Studies have shown that even for first-time instructors, such as the Teaching Assistants leading the review session, Peer Instruction helps students retain material.[4] Peer instruction allows students to come to an understanding of concepts while discussing questions with their classmates. Porter et al. found that students are actually learning as they discuss with each other and not simply copying what someone else believes to be the right answer. Porter et al. measured learning by using “Weighted

Learning Gain”. Weighted learning gain metrics are applied to a series of peer instruction questions during a particular session. Our study uses this methodology.

This project focuses on CS1. However, if successful, a similar procedure can be applied to other classes in the curriculum [4]. Peer Instruction is different from a regular, direct instruction, lecture. Throughout a lesson, the instructor presents a multiple choice question to the students. They answer the question first on their own. They then discuss the question with their peers and answer a second time. The main goal of this discussion is that students who felt confident about their answer could explain and teach their peers, solidifying their understanding. Multiple Peer Instruction questions are given, based on a similar concept in hopes that more students get the answer correct after learning from the first question.

We incorporated some of the research Porter et al. did regarding peer instruction and how they measured learning gains. [4] In their research, they wanted to see the efficacy of peer instruction and see if students were actually learning or just copying from their peers. They were able to define a new metric, Weighted Learning Gain, which better reflects the learning value of group discussion. Peer instruction was used during the review sessions and to more accurately measure the learning gains of the students we used the method created by Porter et al.

1.2.2 Teaching Assistant Led Review Sessions

The new format and content of review sessions are a result of updating work by Gilbert et al.[1] The key difference between what Gilbert et al. accomplished and what this project accomplished was that the prior project was in Java while this one is in Python and incentivized review session attendance with weekly in-class assessments. The sessions were designed using peer instruction

as it is effective with new instructors. [7] Since the original Fourth Hour was created, the CS1 class has been converted from Java to Python and thus new material must be generated. The difficulty of this study is that we want to target students who will benefit from the review sessions and ensure they come without also encouraging too many students who are already doing well as these high performing students can intimidate those who are struggling.

In Fall 2021, two instructors experimented with weekly in-class assessments and offering retakes after review session attendance. Fourth Hour attendance rates improved significantly for those two instructors in Fall 2021. Spring 2022, when these incentives were no longer offered, the attendance dropped to almost zero, though the teaching assistants and faculty in several sections agree that there are still several students who could benefit from the review. Since the start of the Python-based review sessions, with no weekly in-class assessments, there has been a significant drop in attendance. Comparing the number of students who attended review sessions in Fall 2021 to current attendance indicates a link between attendance and in-class assessments.

An alternative but less feasible approach would be to have faculty lead review sessions. Dillon and Slattery [6] revised the curriculum at Michigan State University in an attempt to increase retention of computer science material. They created a one-credit course to supplement the CS1 course. Lack of resources made this infeasible long-term at JMU as this supplemental course was run by professors, not TAs. This is also not applicable to this study as the class size in James Madison University's CS1 course has only 30 students, and its instructors already consistently use active learning in classes and office hours, while Michigan State's course being studied was large (120 students) and primarily online or straight lectures.

Chapter 2: Methodology

In the previous section, we discussed the way others have approached different metrics that we are using in this research. In this section, we explain these metrics in more detail and discuss the specific metrics we used in our implementation.

2.1 Student Attendance

To give the students who attended the review session credit for attending, before each Fourth Hour session began, we had one of the Teaching Assistants leading the session walk around the room and take attendance. They collected three things from each student: their name, their professor's name, and whether they had been to class the previous week. The collected names were later reported to their professor so the students could be given another attempt at their weekly quiz.

2.2 Content Retention

2.2.1 Measuring Learning Gains

To measure learning gains during each session, a similar approach to Porter et al was taken.[4] We used the potential learner group (PLG), students who got the first attempt of question 1 incorrect and after discussing with their peers, got the question correct during the second attempt, and the control group (CG), students who got question 1 correct on both attempts. We then determined if the students in the PLG and CG got the first attempt of question 2 correct. We took these scores of the PLG and divided them by the CG to get the weighted learning gain (WLG):

$$WLG = \frac{PLG \% Correct}{CG \% Correct}$$

By calculating the WLG, we can see if students who attended the review session get value out of attending and then determine if this impacts their grade on examinations and ultimately, the course.

We also computed learning gains for the whole semester from an in-class pre- and post-assessment. In our study, we used the first exam as the pre-assessment and the final exam as the post-assessment, as we felt that this more accurately demonstrated the students' understanding of concepts. The formula below is based off of Gilbert et al. but, we are using the final exam and the first exam rather than a pretest and a posttest. Students with higher learning gains improved more throughout the semester, than those with lower learning gains.

$$\frac{(Final-Exam1)}{(100-Exam1)}$$

2.2.2 Data Analysis

This study took a quantitative approach to determine the effectiveness of each weekly unit. We analyzed the data taken in Fall 2022. By using the results of the peer instruction questions, we observed if the students successfully understood the material described by the learning objectives and if learning happened during the session. This has the advantage of generating data on the session itself, as well as indicating to the TA instructors which weekly concepts students are struggling with the most.

Chapter 3: Implementation

3.1 Study Session Design

3.1.1 Previous Approach

The previous approach gave the students the opportunity to come to a review session three days a week. The sessions were offered at different times to accommodate students' schedules. The content was the same for all three sessions. Two teaching assistants led each review session, one to present and one to support the other. Having two TAs assigned to each review session increases fidelity by ensuring that peer instruction is followed and the presenting TA explains the concepts clearly. The other TA is also there to support the other if there is a complaint or a question that the other TA is unable to resolve themselves. These TAs also alternate who is presenting the material to keep the material fresh and students engaged. In the Spring 2022 semester, students had little incentive to come to these sessions, and the sessions were primarily attended the day before an examination. In the Fall of 2021, two faculty members experimented with weekly in-class assessments and offered retakes to review session attendees, which increased attendance as the students realized they needed more help. This approach to incentivize attendance is what is being studied more formally in a coordinated study for Fall 2022.

3.1.2 Current Approach

Our current approach, and the approach utilized in this study, is to make two peer instruction questions that will be given during the review session. This will encourage students to pay attention to the material being taught in the session so they can get the questions correct. These

questions are not graded and the answers will be anonymous during the session. Students will pay attention so they can contribute to their group during the discussion. Many more students could benefit from attending the session, but they either do not feel they need to or they lack the time. By adding the weekly in-class assessments and retakes for this study, students would learn that they do not actually understand the material as well as they previously thought and then would attend the review session. This study has translated the review sessions to Python for the Fall semester 2022 making sure that peer instruction questions address student misconceptions in the literature and are appropriate for measuring learning during the session.

We used Socrative to administer the peer-instruction questions, allowing for discussion about all answers, both correct and incorrect answers with each student's assigned group. Socrative is a website that allows students to easily answer a question given by the instructor on their personal device. We hope this will give the students a feeling of anonymity, making them more likely to participate. [5] By answering the question in Socrative, students are also committing to an answer that, if incorrect, they will know they did not understand completely.

3.1.3 Session Schedule and Misconceptions Covered

The schedule for the Fourth Hour can be seen in Table 3.1.3. The content of each review session was designed around material covered in CS1 class the previous week. Each week the questions were designed to target common misconceptions found in Kaczmarczyk et al. [2] The misconceptions can be seen in the top row with the weeks and their topic shown in the first column. Most of these misconceptions are self-explanatory. To remain uniform with Gilbert et al. and improve clarity, we refer to "primitive no default", which refers to a student

misunderstanding that instance variables of a class have no default value, as “instance variable no default”.

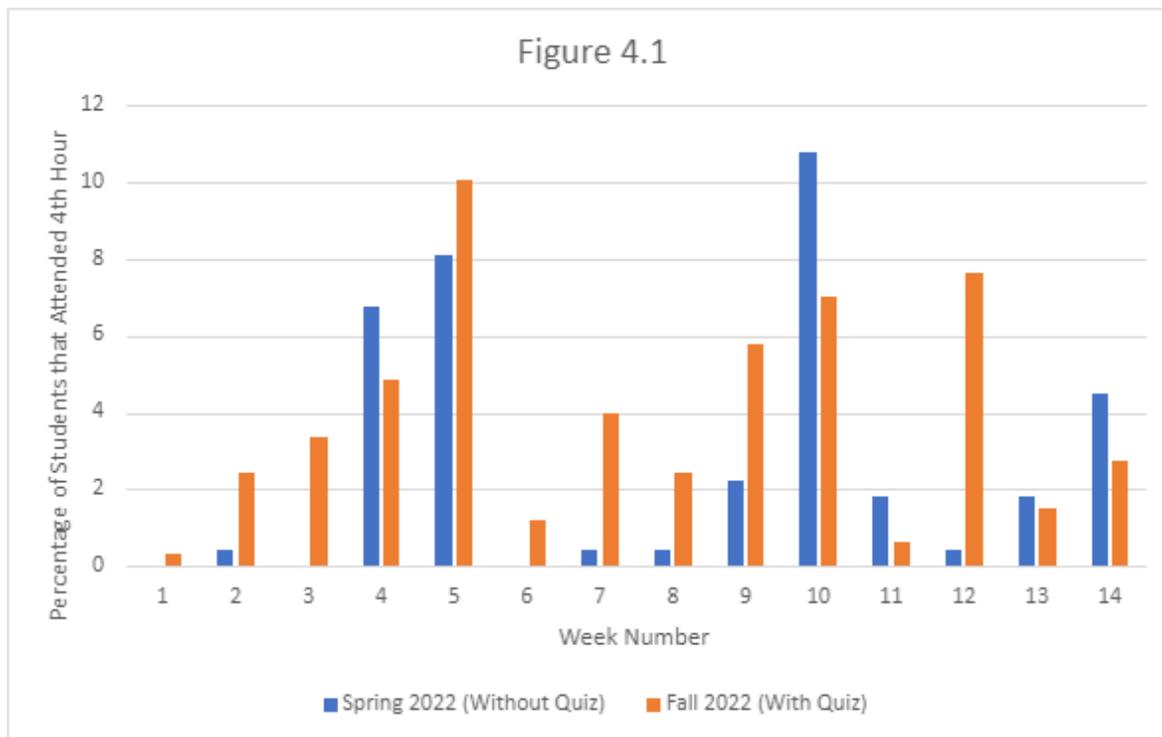
Table 3.1.3

	Language elements & memory usage	While loop operation	The Object concept	Cannot trace code linearly	Semantics to variable declarations	Off by 1 array construction	Instance variable no default
Week 1: Introductions	X						
Week 2: Variables and Operators	X				X		
Week 3: Data Types	X				X	X	
Week 4: Functions	X			X			
Week 5: Exam Review I	X				X		
Week 6: Testing							
Week 7: Conditionals	X			X			
Week 8: While Loops	X	X			X		
Week 9: For Loops	X						
Week 10: Exam Review II	X	X		X	X		
Week 11: Unit Testing							
Week 12: Classes I	X		X		X		X
Week 13: Classes II	X		X		X		X
Week 14: Exam Review III	X	X	X	X	X		X

Chapter 4: Results

4.1 Attendance

In Fall 2022 we introduced a weekly review quiz that incentivized students to attend The Fourth Hour review session. This incentive was not available in previous semesters namely Spring 2022. Using the attendance data we gathered at the start of every review session, we were able to determine the percentage of students that attended The Fourth Hour every week for both Spring and Fall 2022 to compare if the weekly quizzes had an impact on attendance. The percentage of all enrolled students was chosen as the number of enrolled students differed per semester. This scale makes it easy to compare the two semesters. The results shown in Figure 4.1 show a trend of increased attendance in the Fall of 2022 when the review quiz was given and students were incentivized to attend The Fourth Hour. Without the incentive, students seem to only attend the review session before an exam, week 5, 10, and 14. The only other time high attendance is seen is week 4, when we covered functions. Functions can be a difficult topic to understand for many students so that is possibly why we saw more non incentivized students attending that week.



4.2 Learning Gains During The Fourth Hour

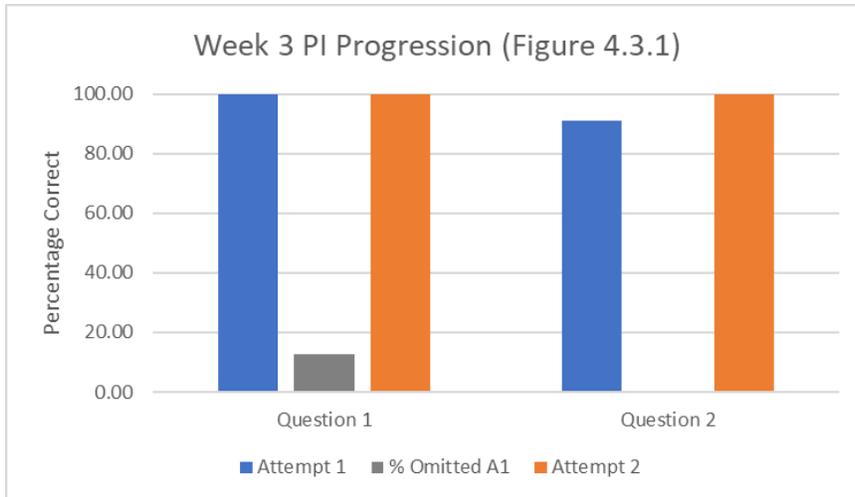
To measure the learning gains during The Fourth Hour review session, we exported the data from Socrative then isolated the Peer Instruction questions. Both Peer Instruction questions target the same misconception with the 2nd question supposedly being more difficult than the first, although similar enough that we had hoped students would be able to make the connection to the first question. This way we could see how many students were able to learn the concept from the first question and apply it to the second question, given later in the review session.

After collecting and analyzing the data we noticed some of the questions may have been too easy or too difficult for the students. It is difficult when creating the Peer Instruction questions to gauge difficulty as we do not know how well the students who show up to The Fourth Hour on any given week will understand the concepts. We hoped the students who struggled on the

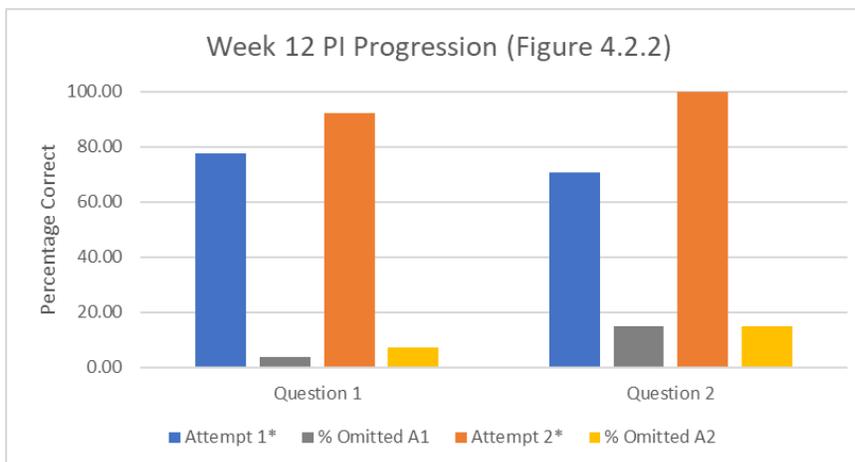
quizzes would be the majority of the students in attendance, but welcome all students who attended since students who attended were at various levels of understanding, writing questions was challenging. With a longer term study we feel strongly we could dial in the question difficulty, it would likely be more effective to use validated peer instruction questions from the peer instruction site and to coordinate with at least one of the CS1 course instructors to accurately target both misconceptions and the learning objectives for the week.

Another issue we encountered was students would fail to answer one of the questions or the attempts. For example, we believe a student who was confident they got Attempt 1 correct would not bother to put the same answer down for Attempt 2 after consulting with their classmates. We also saw students who would not answer Attempt 1, but would answer after talking to their classmates. We believe this is for the opposite reason, the student did not understand the concept at all and potentially felt embarrassed to guess, so they waited and put what their peers selected. Originally we marked no response as incorrect on the graph, but found this did not display the results accurately. We ended up adding a new category to the graphs, when needed, called “% Omitted” and when doing the Percentage correct calculation we did not count students who failed to reply as correct or incorrect. This represented the data much more accurately as seen below in the applicable figures.

For example in Figure 4.3.1, we consider this question too easy as all students got it correct, at least in Question 1. Also ~12% of students did not reply to Attempt 1 so we have a % Omitted A1, representing percentage omitted for Attempt 1.



Another example can be in Figure 4.3.2, where the question successfully targeted the level of the students in attendance and we had to omit more student data. Week 12 was our second most attended review session, and with some students, the number of non-responses increased.



The rest of the week's Peer Instruction Progression graphs can be seen in Appendix A1.

Most weeks we had two Peer Instruction questions, but weeks 1, 2, and 6 were the exception.

With the first week being introductions, we did not have Peer Instruction questions at all as there had only been a single class session, so there was little to review. Week 2 we had four Peer

Instruction questions, this was simply due to a change of plans. Originally we were going to have four Peer Instruction questions, meaning, we would be covering two types of questions. We quickly realized this took up too much time during the review session and adjusted for future weeks. Week 6 only had one Peer Instruction question; this was due to the nature of post exam weeks. We struggled with attendance and content these weeks, as there was no new material taught the previous week and thus, no incentive for students to attend The Fourth Hour. Week 6 we covered testing which mostly included the TA debugging and stepping through code for the attendees to view so this review session was more involved and less conceptual, thus more difficult to assess their learning with online questions. We ended up resolving this for subsequent post exam weeks, as we could discuss Unit Testing which led to a more traditional Fourth Hour session and thus had two Peer Instruction questions.

Table 4.2 shows the Weighted Learning Gain (WLG) for each week of the Fourth Hour.

Table 4.2

Week	WLG	WLG(%)
2	0.5	50%
3	1.111	111%
4	0.833	83.3%
5	DIV/0	DIV/0
6	N/a	N/a
7	0.96	96%
8	DIV/0	DIV/0
9	2.29	229%
10	0	0%

11	DIV/0	DIV/0
12	0.77	77%
13	1	100%
14	0.5	50%

The data above shows that students that do not originally know the content being assessed and taught, students in the PLG, are learning from the review sessions. We only observed one week where 0% of the PLG students learned the concept, according to the formula.

During week 5, 8, and 11, no one in the control group got the first attempt of question 2 correct. This means that originally, they got both attempts of the first question correct, however, after being asked the second question, all of these students got it incorrect. Because of this, we were dividing by 0 when we did the equation, thus giving us a divide by zero or “DIV/0” error. This has been marked as “DIV/0” in the table.

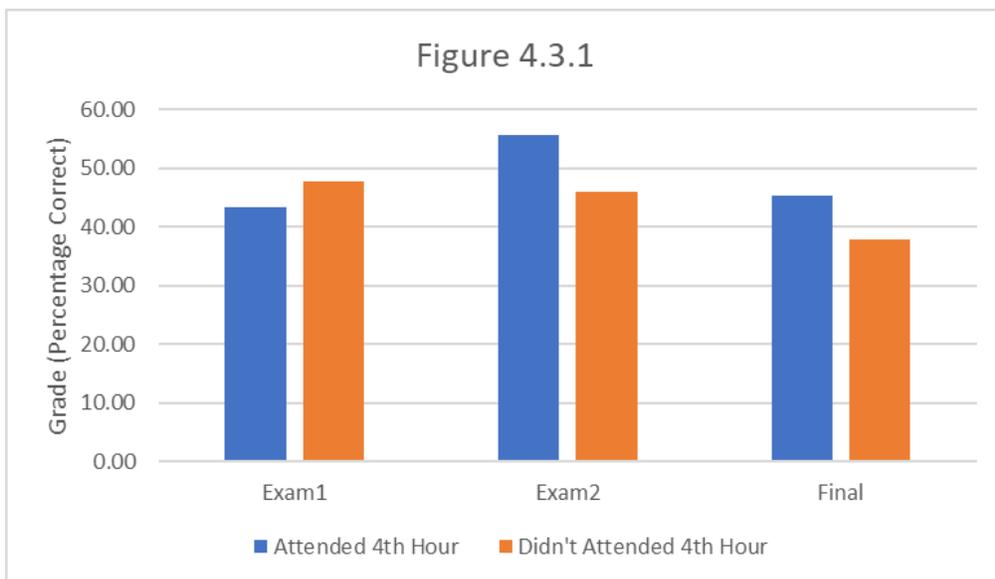
During week 6 we only had one Peer Instruction question, thus could not perform the calculation. This has been marked as “N/a” in the graph.

4.3 Learning Gains Throughout the Semester

Gilbert et al. realized students not attending the Fourth Hour generally had higher learning gains, because many of those students were succeeding in the course. So, we established a subgroup to target, those who did poorly on the first exam who did see learning gain significance. To measure learning gains throughout the semester we first needed to identify a group of students that we

predicted might need additional support. We found that students who received less than an 80% on the first exam would be a good target demographic as they would need to improve on subsequent examinations to pass the class with a grade sufficient to be admitted to the major. The James Madison University Computer Science department requires students to receive a B average in CS1 and CS1.5 in order to be admitted, and our goal is to help students who need additional support get above this threshold.

We began by analyzing this demographic's grades on Exam 1, Exam 2, and the final and separating them based on whether or not they attended The Fourth Hour. As shown in Figure 4.3.1, on the first exam students who attended the Fourth Hour did worse than students who did not; however, on Exam 2 and the final, the students who attended The Fourth Hour scored approximately 10 points higher on average between the two exams. This result suggests that the review session helped these students score higher than other students who did not attend the review sessions.



We found that the average learning gain of the demographic described above that did not attend The Fourth Hour was **-0.1867**, while the average of the students who did attend was **0.0597**. Students with a negative learning gain score received a lower grade on the final than their Exam 1 meaning that on average, students who scored below an 80% on their Exam 1 and did not attend the review session, did not bring their grade up on the final exam, while students who did attend the review session, did bring their grade up.

Chapter 5: Conclusions and Future Work

5.1 Primary Results

In summary, this study found that:

1. When given weekly review quizzes with the opportunity to retake them if the student attended The Fourth Hour, a greater number of students attended the review session.
2. Students who received less than an 80% on Exam 1 and attended The Fourth Hour on average received higher grades on Exam 2 and the Final than students who did not attend The Fourth Hour.
3. Generally, the approach of using peer instruction in a review session taught by TAs was an effective learning mechanism for attending students as long as the peer instruction question was well aligned.

5.2 Future Work

Finally, we have some lessons learned that could improve future work.

1. We need to ensure that students answer all of the Socratic questions so we do not need to omit students and have data that more accurately depicts student comprehension of the topic. We suggest that students will not get credit for attending the review session unless they answer all of the Socratic Questions. They would not need to get all questions correct, just make a good faith attempt to receive the quiz retake. This approach will also ensure students are trying to learn in the review session and not just sitting in the back on their laptop or phone.

2. Each week peer instruction questions should be reviewed to make sure they are targeting the right level and aligned with the learning goals and misconceptions of the specific course sequence that semester. This is relatively easy as not every question needs to be improved. Simply look at the Socrative data from the previous semester and see if the scores were too high or low and adjust the difficulty accordingly.

Bibliography

- [1] Gilbert, Megan Elizabeth, "A pilot study on the impact of teaching assistant-led CS1 study sessions using Peer Instruction" (2020). Senior Honors Projects, 2020-current. 28.
<https://commons.lib.jmu.edu/honors202029/28>
- [2] Kaczmarczyk, Lisa & Petrick, Elizabeth & East, J. & Herman, Geoffrey. (2010). Identifying student misconceptions of programming. SIGCSE'10 - Proceedings of the 41st ACM Technical Symposium on Computer Science Education. 107-111. 10.1145/1734263.1734299.
- [3] Bureau of Labor Statistics, U.S. Department of Labor, Occupational Outlook Handbook, Computer and Information Research Scientists, at
<https://www.bls.gov/ooh/computer-and-information-technology/computer-and-information-research-scientists.htm> (visited March 27, 2022).
- [4] Leo Porter, Cynthia Bailey Lee, Beth Simon, and Daniel Zingaro. 2011. Peer instruction: do students really learn from peer discussion in computing? In *Proceedings of the seventh international workshop on Computing education research (ICER '11)*. Association for Computing Machinery, New York, NY, USA, 45–52.
DOI:<https://doi.org/10.1145/2016911.2016923>
- [5] Socrative, 15 April 2021, <https://www.socrative.com/>.
- [6] Laura Kay Dillon and Michelle Slattery. "A Flipped Active-learning Class to Support Diverse Students in a Large Introduction to Programming Class". 2018 ASEE Annual Conference &

Exposition , Salt Lake City, Utah, 2018, June. ASEE Conferences, 2018.

<https://peer.asee.org/29677> Internet. 08 Apr, 2020.

[7] Cynthia Taylor, Jaime Spacco, David P. Bunde, Andrew Petersen, Soohyun Nam Liao, and Leo Porter, “A Multi-Institution Exploration of Peer Instruction in Practice,” Proceedings of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE 2018, 2018, doi:10.1145/3197091.3197144

[8] Cory Longenecker. “Analyzing Motivation and Self of Belonging in CS1 Review Sessions”
JMU Scholarly Commons

Appendix

A.1 Peer Instruction Progression Charts:

