HUMANIZING ORGANIC CHEMISTRY

by

ANAIS LILY KOLESNIKOV

A THESIS

Presented to the Department of Chemistry and Biochemistry
and the Robert D. Clark Honors College
in partial fulfillment of the requirements for the degree of
Bachelor of Science

May 2024

An Abstract of the Thesis of

Anais Kolesnikov for the degree of Bachelor of Science
in the Department of Human Physiology to be taken June 2024

Title: Humanizing Organic Chemistry

Approved: *Adam Glass, Ph.D.*
Primary Thesis Advisor

This thesis embarks on an epic voyage to explore the perception of organic chemistry as a complex and abstract subject in the eyes of students, exploring methods to humanize the discipline by enhancing engagement and comprehension in applicable and sustainable ways. The study addresses what makes a course accessible, the value of interdisciplinary approaches to STEM, the benefits of incorporating humanities techniques into STEM education, and the impact of active learning on organic chemistry outcomes.

The procedures employed in this study involve two main facets: course improvement and analysis. Course improvements implemented in the University of Oregon's CH331 were not a solitary effort, but a collaborative one. They include the establishment of a Peer-Led Teaching Team (PLTT), the creation of Active Learning Workshop Sessions (ALWS), and the administration of Mock Exams (ME). The PLTT, for instance, consists of students who previously excelled in CH331, assisting current students through structured workshops and mock exams designed to reinforce understanding and application of organic chemistry concepts, while establishing a safe learning environment.

The major results of the study reaffirm the benefits of these interventions. They significantly enhance student engagement, understanding, and performance in organic chemistry. Active learning strategies, coupled with peer-led support, foster a more interactive and supportive classroom environment, mitigating the anxiety and confusion traditionally associated with the subject. This reassurance is based on the positive feedback and improved outcomes observed in our study.

The conclusions drawn from this study underscore the transformative potential of rethinking the structure and delivery of organic chemistry courses. By emphasizing critical thinking, collaboration, and real-world applications, educators can make organic chemistry more accessible and enjoyable, ultimately preparing students more effectively for their future academic and professional endeavors. This optimistic outlook is based on the promising results of our interventions, which have significantly enhanced student engagement, understanding, and performance in organic chemistry.

Acknowledgements

First and foremost, I would like to thank Dr. Adam Glass for believing in our vision and trusting the process. The scope and impact that this project has had in the lives of students was only possible with his support and encouragement. He pushed me to go above and beyond, expanding the horizons of our project, evoking the real, tangible changes we witness before us today. I want to thank Dr. Barabara Mossberg, who cheered me on, encouraging me to find my voice and use it to make an immortalizing difference in the world. I wish to thank Dr. Amy Connolly who revolutionized the way I perceive teaching, leading by example and building community and camaraderie at all levels of STEM. Of course, I want to thank all the wonderful Human Physiology, Chemistry, and Biology professors at the University of Oregon who built the safe learning environment we enjoy today. As well, I wish to express my deepest gratitude and appreciation to my parents, my brother, and my friends who supported me, inspired me, and listened to me churn these ideas over and over and… over again until they were perfect.

.

Table of Contents

List of Figures

List of Tables

Introduction

Organic chemistry is not a young science, temporally speaking, first monikered in 1807 by one Jöns Jacob Berzelius. This specific branch of chemistry tells the story of carbon. As the backbone of the natural world, an understanding of the way it forms bonds and interacts with other atoms can be used to explain everything from the viability of an embryo to the scale of destruction caused by a C4 explosion. However, the beauty of this discipline is rarely reflected by the reactions it receives when its name is referenced. The growth of this discipline has been undisputably dynamic, yet its reputation remains stagnant. The publication of Morrison and Boyd's Organic Chemistry textbook in 1959 earned the discipline a mainstream place in pre-med and health-related fields' curricula (Cooper et. Al). Yet, here we are, over 60 years later, and it seems little has changed about the way many students perceive the subject. Age-old critiques still roll off contemporary tongues: organic chemistry is confusing, abstract, sterile, and –the dreaded— "I don't understand why med schools require it". Anatomy? Absolutely. General Chemistry? Of course. Biology? Sure. But what is "Ochem's" role in determining who medical schools scoop up and bestow their greatest gifts upon?

Biology does not get the bad rep that Organic Chemistry sports. Students don't often switch majors or career-paths from a single term of Bio. That seems to be a power reserved for Orgo, as it sits on its mystic stir-plate, swirling in rumors and mystery, titrating the tears of students past, present, and future. Let us address the charging elephant in the room: why is Organic Chemistry so widely despised?

Rumor has it that success in Organic Chemistry can only be achieved through rote memorization, more rote memorization and maybe a prayer. Famed as the weeder course to end all weeder courses, Organic Chemistry has long haunted students and professors nationwide with failure and withdrawal rates teetering on 50% nationwide (Grove et al.) Shockingly, these are pre-pandemic numbers. Test scores across disciplines and across the country are still struggling to recover from the effects of the lockdown. However, the global pause created a space to think. Systems we previously accepted became increasingly futile, the cracks in the foundation of all that we knew about education and how to improve it began to show. Students do not just seem disengaged from the content but from the idea of teamwork itself. As the seasoned academics for whom this project is intended, we know that teamwork makes the dream work. Perhaps it's the general structure of lecture or the post-quarantine social anxiety, but the silence between students is deafening. Asking the professor a question in front of the whole class is not for the faint of heart, just thinking about the prospect of such a feat may tease out a fight-or-flight response, beads of sweat pearling on palms and foreheads. Yet, we know that student engagement is a vital component of learning. Bloom's taxonomy of learning (*Bloom's Taxonomy - Center for Instructional Technology and Training - University of Florida*, 2024). This model explores the connection between the different levels of knowledge acquisition and presents a theory for the importance of each step in the education process. As we can see, understanding and remembering fall on into the categories of lower order thinking skills. The higher order thinking skills include analyzing, evaluating, and most importantly creating. The whetstone that sharpens this set of skills is critical thinking. Critical thinking is achieved through critical understanding. Understanding the "why" is perhaps the most important part of being successful in any field; purpose is our driving force.

Why Organic Chemistry? Organic chemistry is widely recognized for its unique ability to facilitate critical thinking in seemingly abstract circumstances, making it applicable to most endeavors. It is meant to build upon the scientific repertoire acquired through general chemistry courses, imploring students to apply their knowledge using established patterns and relationships in novel situations. Appreciating an introductory, non-majors course such as CH331 Organic Chemistry stems from understanding its role in developing an understanding of connectivity, rationalizing and critical thinking and coming to an educated guess on how to solve that problem. Organic chemistry teaches us to rely on our values, both academically and colloquially. Some of this gets lost in translation, students have expressed the sentiment that key component of Organic Chemistry is building endurance and mental resilience, useful skills yet they fall short at bringing out all the beautiful complexities that organic chemistry has to offer. Bridging the gap between the intended purpose of a course and the way it is perceived through the eyes of students is achieved by defining the goals associated with learning and cultivating the environment conducive to meeting the goals set by the instructor. This is only possible when evaluating and understanding the students' experience and goals.

According to authors Eric M. Anderman and Tamera B. Murdock students typically fall into two goal orientations: mastery and performance (Anderman et Murdock, 2007). Mastery-goal oriented learning, just as it sounds, is when a student cares about learning and *mastering* the content. A performance-goal-oriented student focuses on the final grade, reflecting their performance in the course. Just as sports legends around the world would encourage, "Win, lose, or draw -—have fun", we should remind students that their role in the classroom is to make the most of their time together and the resources they have available. Skills such as communication, teambuilding, and cooperation hold real-world value; these are the traits that make a candidate desirable whether applying for graduate school, medical school, or a job in their chosen field. Organic Chemistry 331 at the University of Oregon has the potential to inspire generations of students to come, by allowing them to bring their humanity into the classroom.

Today more than ever before, we have access to information. Tutorials on how to approach mechanisms and remember SN_1 versus SN_2 reactions have sprouted up like flowers after rain. However, achieving a mastery goal in the realm of Organic Chemistry comes from applying the material, something many students struggle to do on exams. How can educators address this disparity, translating understanding to application? Perhaps the answer lays in the course structure, rather than its content.

Research Questions:

1.) What makes a course "accessible"?

2.) What is the value of an interdisciplinary approach to STEM?

3.) How can techniques taught in humanities courses benefit STEM?

4.) Does active learning benefit Organic Chemistry 331 outcomes?

Methods:

This project is made of two distinct, yet hopelessly interconnected facets: improvement and analysis. The changes to the course itself fall under the category of course improvement, quality improvement as defined by the IRB. The changes made to CH331 are not driven by research; instead, research was used to identify ways to improve the course then implement the changes for the entire course's betterment. The changes and their implementation will be described in the following sections, separate from the process of obtaining data and its statistical analysis.

CH331 Course Improvement and Implementation:

There are three primary components to be described functioning as part of the general CH331 course improvement: the establishment of a Peer-Led Teaching Team, creation and implementation of Active Learning Workshop Sessions, and Mock Exams.

- Establishing a Peer-Led Teaching Team (PLTT): The peer-led teaching team is headed by the instructor and Graduate Employee (GE), comprised mostly of Peer Learning Assistants (PLAs). The PLAs were students who had previously taken and excelled in the CH 331 course. Weekly meeting sessions were held to discuss weekly goals and intended learning outcomes for students. During these meetings, peer leaders were encouraged to work through the worksheets students would be asked to do the following Friday, identify where problems may arise, and develop "elevator speeches" to summarize key concepts. These meetings were a chance for the instructor to correct the worksheets and any misconceptions expressed by the teaching team, serving as a safe space where leaders could ask questions about each topic. Weekly meetings were necessary to ensure students were getting accurate information from the PLAs and that the PLAs had practice explaining the concepts in an accessible way. Peer learners were also encouraged to work in small groups, analogous to the Active Learning Workshop Sessions, which allowed them to practice both teaching and approaching students. (prepared, cohesive, INSPIRE model)

- Active Learning Workshop Sessions (ALWS): These workshops were designed by our teaching team. Most sessions focused on working through the worksheet provided and other sessions (those preceding exam weeks) featured mock exams (described below). Workshop sessions were held every Friday, each section was allotted 2 hours. Attendance was taken once the worksheet was completed and students could leave as they finished or use the remaining time to ask the teaching team questions about the content.
- Mock Exams (ME): During the sessions that directly preceded exam weeks, we incorporated mock exams with specific guidelines during ALWS. Students were provided with sample exam questions and asked to work in their table groups to come up with the correct answer. Students were then brought together by the professor or GE and asked to discuss the questions. Students were not provided with the answer in the moment but instead encouraged to discuss through Socratic-style guidance until the correct answer was reached. Certain barriers had to be overcome such as the reluctancy to answer in front of the whole class. This issue was resolved when students were offered a single extra credit point for their participation. This is not enough to significantly impact the student's scores but proved to be enough to ignite discourse, engaging students to participate. The other component that helped facilitate discussion was the soft transition from instructor-led sessions to PLA-led sessions. We observed that students seemed more willing to participate, perhaps empathizing with the position of the peer-leader,

volunteering more readily to answer a question to avoid putting their peer in an awkward position (i.e. silent classroom). Perhaps students were also less anxious to make a mistake if it was in front of student leaders as opposed to professors. Cultivating the proper environment for the workshops was crucial to the success of these sessions. Identifying problems and working to fix them in the moment allowed us to tailor the ALWS to our specific population of students, maintaining support and encouragement throughout the process.

Course Comparison:

Data was acquired from three years of CH331: 2021, 2022, 2023. Each year, the course was taught by Dr. Adam Glass. The structures of the courses are as follows:

- ’21: standard lecture 4 days a week (Monday-Thursday) 50 minutes each
- ’22: standard lecture 3 days a week (Monday, Wednesday, Friday) 80 minutes each, with active learning ~15 minutes built into the lecture, iClicker questions and encouragement to collaborate with neighboring colleagues.
- ’23: Standard lecture 3 days a week (Monday, Wednesday, Friday) 50 minutes each, Mandatory 2-hour active learning workshop sessions (ALWS), mock exams every ALWS that fell on Fridays right before exam date.

IRB Approval and Exemption:

IRB Exemption was granted for this project on (2/6/2024). Please see attached appendices for official statement.

Statistical Analysis:

For statistical analysis in this project, we performed the Kruskal-Wallis Test when comparing three independent samples across the three years of interest. We also performed the Wilcoxon Rank Sum Test when comparing data from 2022 and 2023. We chose these tests as they compare the means and standard deviations of independent samples and do not assume certain parametric dispositions when assigning ranks, characteristic of our studied population.

Results:

Characteristic	N	2021, N = 307[1]	2022, N = 241[1]	2023, N = 293[1]	p-value[2]
Exam_1	834	67.00 (54.00, 80.75)	77.00 (67.00, 86.00)	79.00 (64.00, 89.00)	<0.001
(Missing)		1	4	2	
Exam_2	831	68.00 (54.00, 79.50)	78.00 (64.00, 89.00)	77.00 (64.00, 86.00)	<0.001
(Missing)		0	4	6	
Exam_3	819	75.00 (57.00, 88.00)	78.00 (66.00, 88.00)	82.00 (66.00, 90.00)	0.002
(Missing)		2	8	12	
Final	815	62.00 (50.67, 75.67)	71.00 (57.00, 84.00)	91.00 (70.50, 106.50)	<0.001
(Missing)		4	8	14	

[1]Median (IQR)

[2]Kruskal-Wallis rank sum test

Table 1: Average exam scores across all three conditions (2021: traditional lecture; 2022: traditional lecture with in-class active learning; 2023: traditional lecture, in-class active learning, and weekly 2-hour ALWS) with Q1 and Q3 in parentheses. N represents number of scores sampled for each exam across the three conditions. Scores across conditions in the same exam were compared using the Kruskal-Wallis Test. P-values shown in the final column.

Table 1 demonstrates the median scores with the interquartile range (IQR) in parentheses. The information above supports the distributions observed across all conditions and exams in the following figures. P-values are found using the Kruskal-Wallis Test, and all were found to be significant.

Characteristic	N	2021, N = 3071	2022, N = 2411	2023, N = 2931	p-value2
Exam_1	834	66.42 (18.01)	75.08 (15.38)	75.99 (15.71)	<0.001
(Missing)		1	4	2	
Exam_2	831	65.95 (17.41)	74.81 (17.84)	73.67 (16.71)	<0.001
(Missing)		0	4	6	
Exam_3	819	70.31 (21.94)	75.87 (16.50)	76.63 (17.73)	0.002
(Missing)		2	8	12	
Final	815	61.55 (19.17)	70.01 (19.91)	87.41 (24.43)	<0.001
(Missing)		4	8	14	

1Mean (SD)

Table 2: Average exam scores across all three conditions (2021: traditional lecture; 2022: traditional lecture with in-class active learning; 2023: traditional lecture, in-class active learning, and weekly 2-hour ALWS) standard deviations in parentheses. N represents number of scores sampled for each exam across the three conditions. Scores across conditions in the same exam were compared using the Kruskal-Wallis Test. P-values shown in the final column.

Table 2 demonstrates the mean scores with the standard deviation in parentheses. The information above supports the distributions observed across all conditions and exams in the following figures. P-values are found using the Kruskal-Wallis Test, and all were found to be significant. It is important to note that the mean value is best understood in the context of the distributions, as it can be affected by skew. The mean value is more representative of the scores students received when the data distribution is closer to being symmetrical.

Characteristic	N	2022, N = 241[1]	2023, N = 293[1]	p-value[2]
Exam_1	528	75.08 (15.38)	75.99 (15.71)	0.4
(Missing)		4	2	
Exam_2	524	74.81 (17.84)	73.67 (16.71)	0.3
(Missing)		4	6	
Exam_3	514	75.87 (16.50)	76.63 (17.73)	0.2
(Missing)		8	12	
Final	512	70.01 (19.91)	87.41 (24.43)	<0.001
(Missing)		8	14	

[1]Mean (SD)

[2]Wilcoxon rank sum test

Table 3: Average exam scores across the two featured conditions [2022: traditional lecture with in-class active learning & 2023: traditional lecture, in-class active learning, and weekly 2-hour ALWS] with standard deviations in parentheses. N represents number of scores sampled for each exam across the three conditions. The “Missing” Column shows the scores not recorded (students did take that exam or receive score). Scores across conditions in the same exam were compared using the Wilcoxon Rank Sum Test, with p-values shown in the final column.

Table 3 is designed to show the relationship in mean scores and standard deviations between the 2022 and 2023 cohorts. This allows us to utilize the Wilcoxon rank sum test as we are comparing two conditions. 2021’s cohort was undoubtably affected by the COVID 19 pandemic, and students might have faced extra stressors such as transitioning back to in-person lecture for the first time since the lock down. The choice was made to compare these two conditions as they might be less impacted by extraneous factors such as COVID 19, allowing us to see a more detailed relationship between scores across different conditions and over time. We see that the observed p-value steadily decreases over time yet does not yield significance until the Final Exam. This suggests that the ALWS curriculum works best in series and demonstrates student improvement over the course of the term.

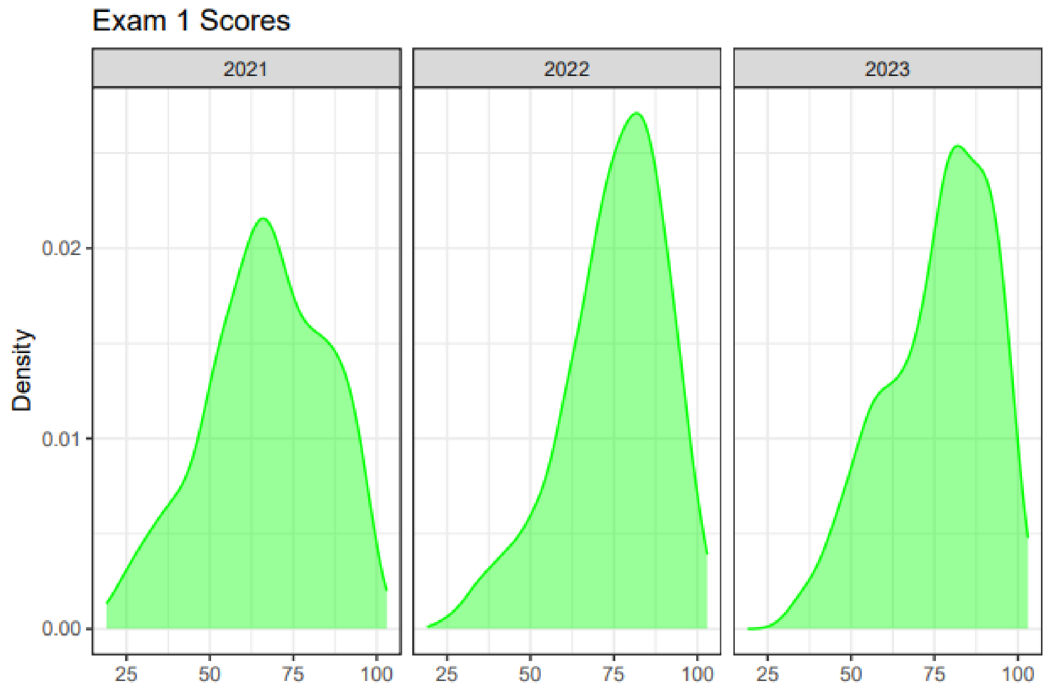

Figure 1 : Exam 1 score distributions across the all three conditions [2021: traditional lecture; 2022: traditional lecture with in-class active learning; 2023: traditional lecture, in-class active learning, and weekly 2-hour ALWS] plotted on x-axis. Density is shown on y-axis.

Figure 1 demonstrates the scores over the three observed conditions. We see here that visually 2021 has the greatest degree of variability, suggesting that student's exam scores were the least uniform across the board. This is further supported by the findings of Table 2, in which the 2021 condition had the greatest standard deviation in scores, found to be 18.01 points compared to 15.38 points and 15.71 points in 2022 and 2023 respectively.

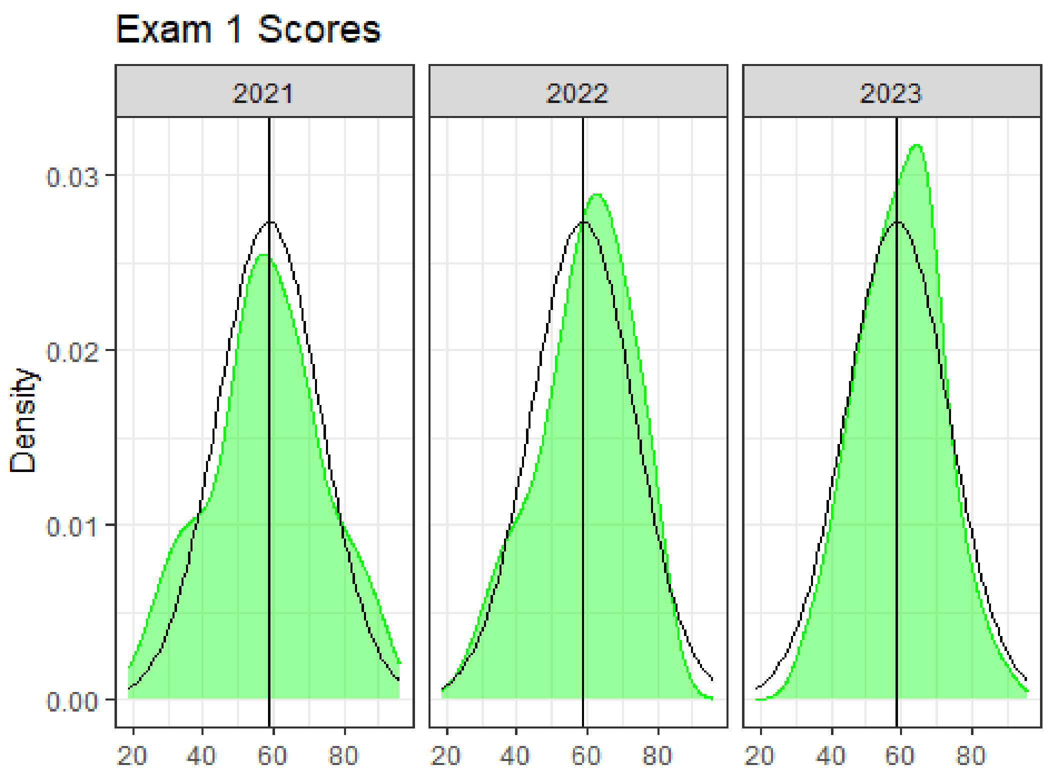

Figure 2 : Exam 1 score distributions across the all three conditions [2021: traditional lecture; 2022: traditional lecture with in-class active learning; 2023: traditional lecture, in-class active learning, and weekly 2-hour ALWS] plotted on x-axis. Density is shown on y-axis. Normal distribution curve is shown in black.

The normal distribution curve shown in black allows us to visualize the variability of the data more clearly. As previously mentioned regarding Figure 1, Figure 2 supports the conclusion that the greatest variability in scores can be observed in 2021. The superimposed normal curve also allows us to see that the variability is skewed to the left of the median, suggesting that the number of students receiving scores lower than the predicted range was greater than those receiving scores higher than the predicted range. Between 2022 and 2023, the normal distribution curve allows us to see that the mode falls close to the median and mean scores, being more pronounced in under the 2023 condition, suggesting that the standard deviation between scores is decreasing as well as that mean scores are increasing. This is further supported by Table

2, in which we can see that the mean score from 2021 increases 8.66 points to 75.08points in 2022 and 9.57 points to 75.99 points in 2023. By comparing the change between just these two scores using the Wilcoxon Rank sum test as shown in Table 3, we do not see a p-value of statistical significance, yet. The p-value is shown to decrease with each subsequent exam, reaching a statistically significant difference by the Final Exam, suggesting that the change was driven by administering ALWS sessions consistently throughout the course.

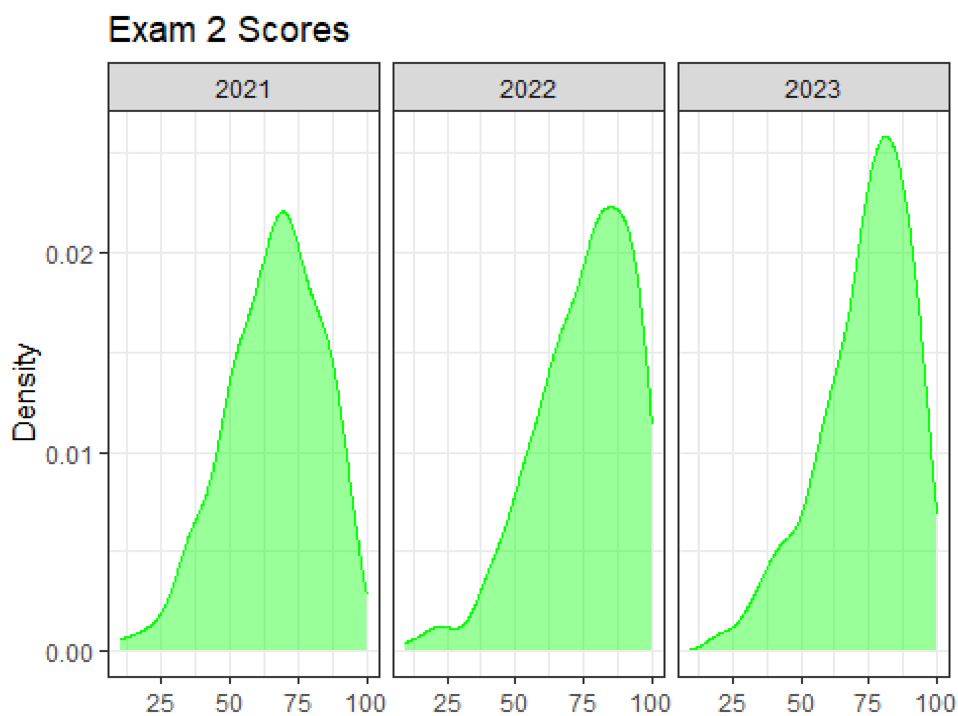

Figure 3: Exam 2 score distributions across the all three conditions [2021: traditional lecture; 2022: traditional lecture with in-class active learning; 2023: traditional lecture, in-class active learning, and weekly 2-hour ALWS] plotted on x-axis. Density is shown on y-axis. Normal distribution curve is shown in black.

Figure 3 shows, again, that the 2021 condition had the greatest degree of variability, which is supported by the standard deviations in Table 2. However, the Exam 2 proved to have generally lower standard deviations across the conditions than any other exam category. The mean scores for both 2022 and 2023 were significantly higher than 2021 yet did not differ much from each other statistically. Visually, we can see that both the 2022 and 2023 conditions have a greater negative skew, as most of the data points fall above the mean scores.

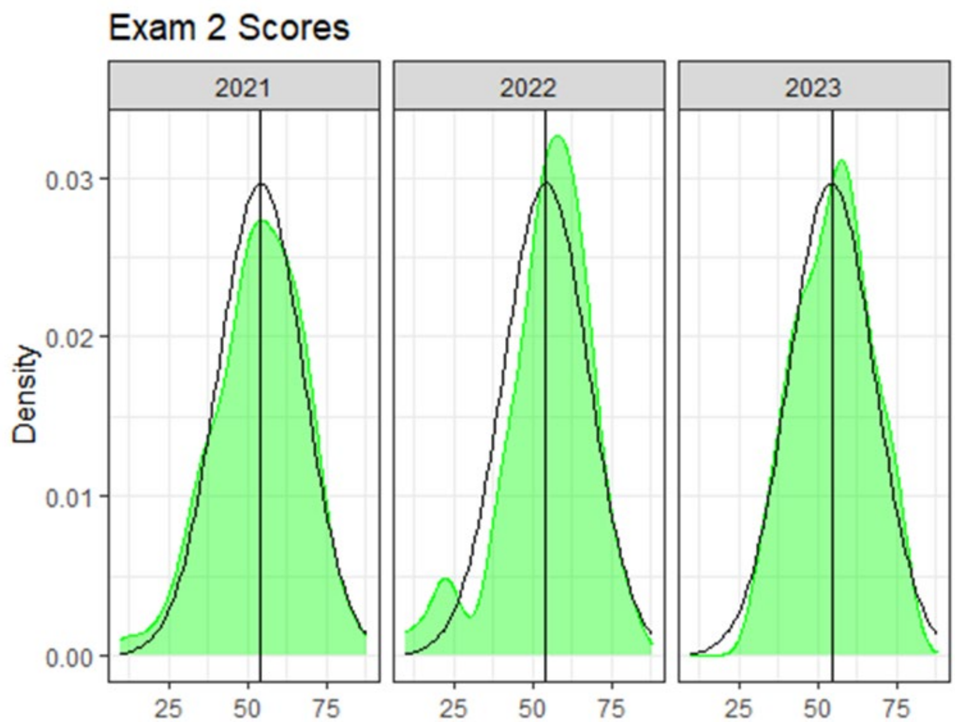

Figure 4 : Exam 2 score distributions across the all three conditions [2021: traditional lecture; 2022: traditional lecture with in-class active learning; 2023: traditional lecture, in-class active learning, and weekly 2-hour ALWS] plotted on x-axis. Density is shown on y-axis. Normal distribution curve is shown in black.

Figure 4 helps us further visualize the data's relationship to a normal distribution. 2022 is one of the more notable distributions in this category as there is a non-symmetric bimodal pattern in the data. This bimodal distribution may indicate a significant group of students struggling to understand apply the material, which is reflected again in Figure 7, under the same 2022 condition. The 2021 condition's distribution shows that the normal distribution curve is an overestimation of the mode of the data, suggesting that there is more variation in the data than anticipated, despite the decrease in standard deviation of this specific category. In fact, Table 2 shows that Exam 2 had the lowest standard deviation for the 2021 condition. 2023's distribution matches most closely the normal distribution, retaining a slightly negative skew of -0.598 compared to -0.57 observed with the previous exam, under the same condition. Less scores are observed on either tail, which is supported by this condition having the lowest variance across the Exam 2 category. This is further supported numerically by Table 2, with the standard deviation of this condition being the lowest, 16.71 points, and the other conditions having 17.41 points and 17.84 points for ’21 and ’22 respectively.

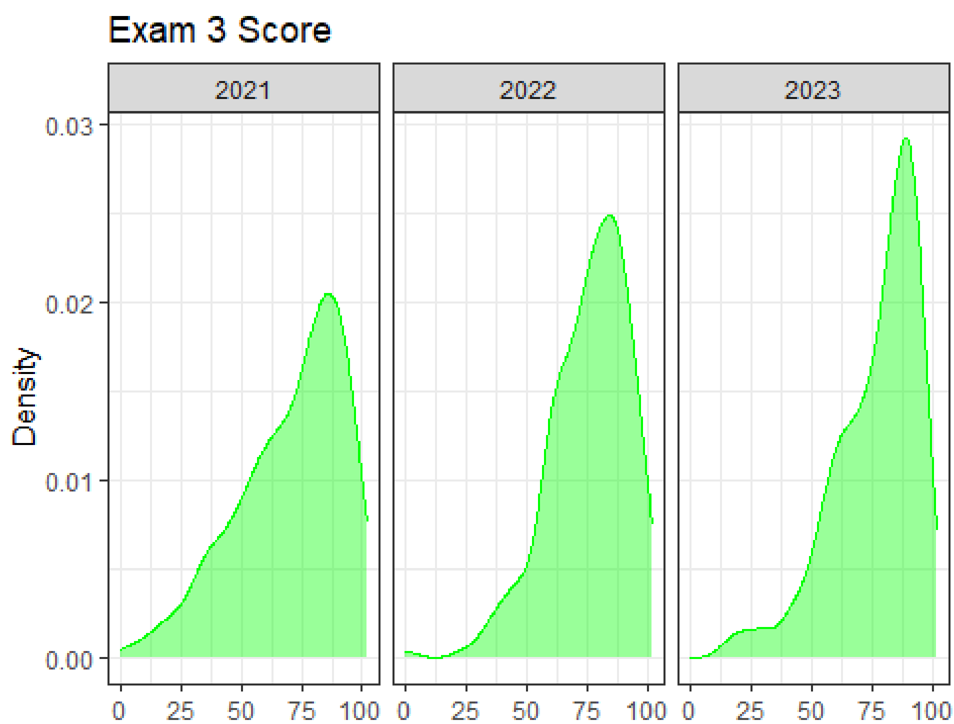

Figure 5 : Exam 3 score distributions across the all three conditions [2021: traditional lecture; 2022: traditional lecture with in-class active learning; 2023: traditional lecture, in-class active learning, and weekly 2-hour ALWS] plotted on x-axis. Density is shown on y-axis.

Each of the conditions across the Exam 3 category show a negative skew, with the modes of the distributions being higher than the means and median scores. To calculate skew that is less influenced by outliers or extreme value(https://www.geeksforgeeks.org/coefficient-of-skewness/), we can use Pearson's Second Coefficient Formula. Shown below is Pearson's Second Coefficient of Skewness Formula:

Pearson's Second Coefficient = 3(Mean − Median)/Standard Deviation

Also written as:

$$Sk_2 = \frac{3(\bar{X} - Md)}{s}$$

Where:
$\bar{X}$: Mean
Md: Median
s: Standard Deviation

Calculating the coefficient of skewness for each of the following conditions can give us more insight into what the shape of the distribution means and whether the Gaussian curve is a good fit for the data. The coefficient of skewness for each of the observed conditions is -0.64, -0.39, and -0.91 for 2021, 2022, and 2023 respectively. Being that each of these values are within the +2, -2 range, we can conclude that the Gaussian curve is a decent fit for the data. Furthermore, we can assert that with a coefficient of -0.39, 2022 data exhibits approximate symmetry, as it falls between the -0.5 and 0.5 range, whereas 2021 and 2023 are both moderately skewed, falling into the -1 and -0.5 range.

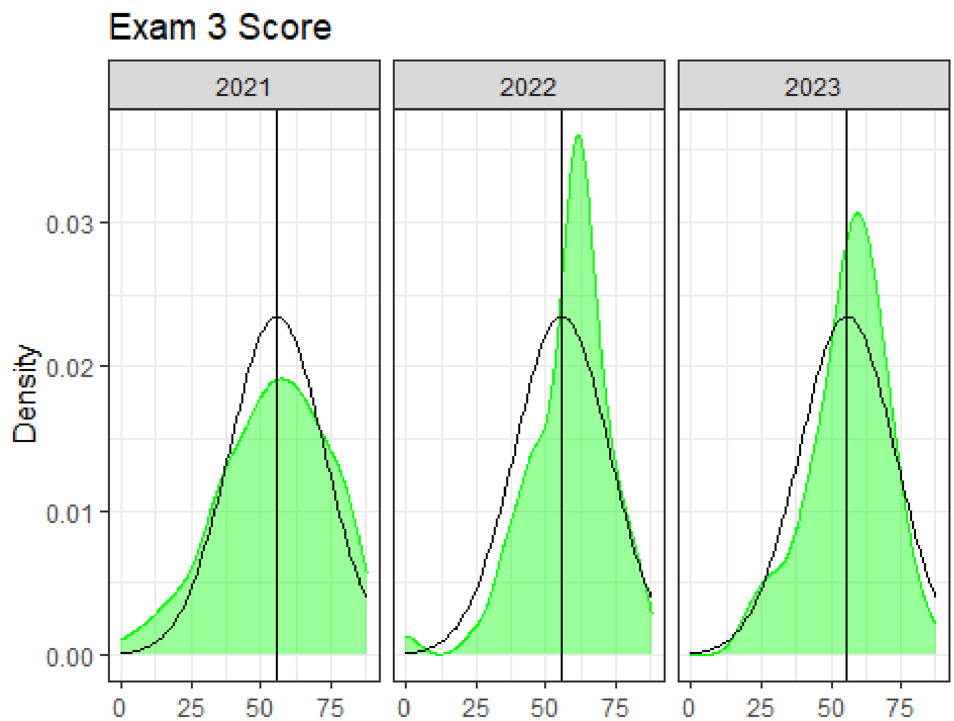

Figure 6 : Exam 3 score distributions across the all three conditions [2021: traditional lecture; 2022: traditional lecture with in-class active learning; 2023: traditional lecture, in-class active learning, and weekly 2-hour ALWS] plotted on x-axis. Density is shown on y-axis. Normal distribution curve is shown in black.

By superimposing the normal distribution curve on the Exam 3 data, we can see that there are dramatic visual changes in the Kurtosis for each of the observed conditions. Kurtosis is defined as the statistical measure that described the shape of a probability distribution's tails relative to is peak(https://www.analyticsvidhya.com/blog/2021/05/shape-of-data-skewness-and-kurtosis/). Often used in financial analysis situations to mitigate risk for an investment, it can be applied to realm of education where we seek to support as many students as possible. The degree to which data values are concentrated around the mean helps us understand student performance in a broader context. Here, we can see that the 2021 cohort has a platykurtic distribution, with lighter tails (meaning outliers occur less frequently) and flatter peak than the superimposed Gaussian curve. For 2021, we do not see as many students as we expect performing at the level of the mean. This observation is further supported by Table 1, as the 2021 condition holds the greatest range value among all three conditions with a Q1 of 57 points and a Q3 of 88 points.

This dispersion suggests the mean itself is not an accurate representation of the average scores of students but rather of the distribution of data itself. The average scores of students in this condition are most likely significantly greater or significantly less than the recorded mean. In the 2022 and 2023 conditions, we see that the peaks are more likely to reflect a leptokurtic trend, in which the observed peaks are greater than the normal distribution. This finding demonstrates that more students scored closer to and above the mean, with a smaller interquartile range among both conditions, indicating less dispersion in the data points. Table 1 supports this finding, Q1 for both 2022 and 2023 was 66 points, significantly higher than that of 2021, and the Q3 was 88 points and 90 points, respectively. Between the two conditions, Table 3 shows that the p-value yielded from the comparison of the 2022 and 2023 conditions was 0.2, not yet statistically significant yet considerably lower than the p-value yielded when comparing the results of Exam 1 (p-value = 0.4).

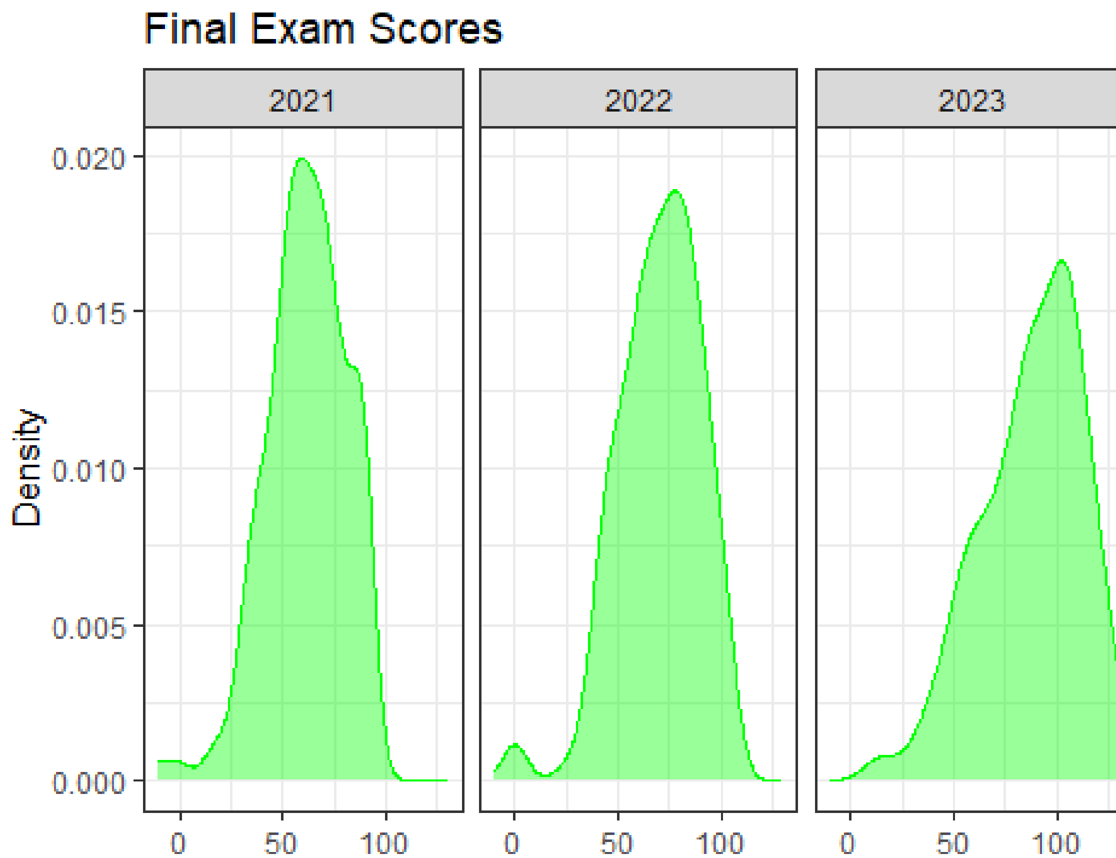

Figure 7: Final Exam score distributions across the all three conditions [2021: traditional lecture; 2022: traditional lecture with in-class active learning; 2023: traditional lecture, in-class active learning, and weekly 2-hour ALWS] plotted on x-axis. Density is shown on y-axis.

Perhaps the most telling of student performance, the Final Exam's cumulative nature allows us to track student learning over time. Qualitatively, the distributions of this exam across all the observed conditions hold the most variation. Upon first look, we can see that the value around which the peak is centered of the 2023 condition is the greatest. This is further supported by the findings of Tables 1 and 2, with 2023 cohort for the Final Exam having a median score of 91 points and a mean score of 87.41 points. For 2022, the median and mean scores are 71 points and 70.01 points, respectively. The median and mean scores for the 2021 condition are 62 points and 61.55 points, respectively.

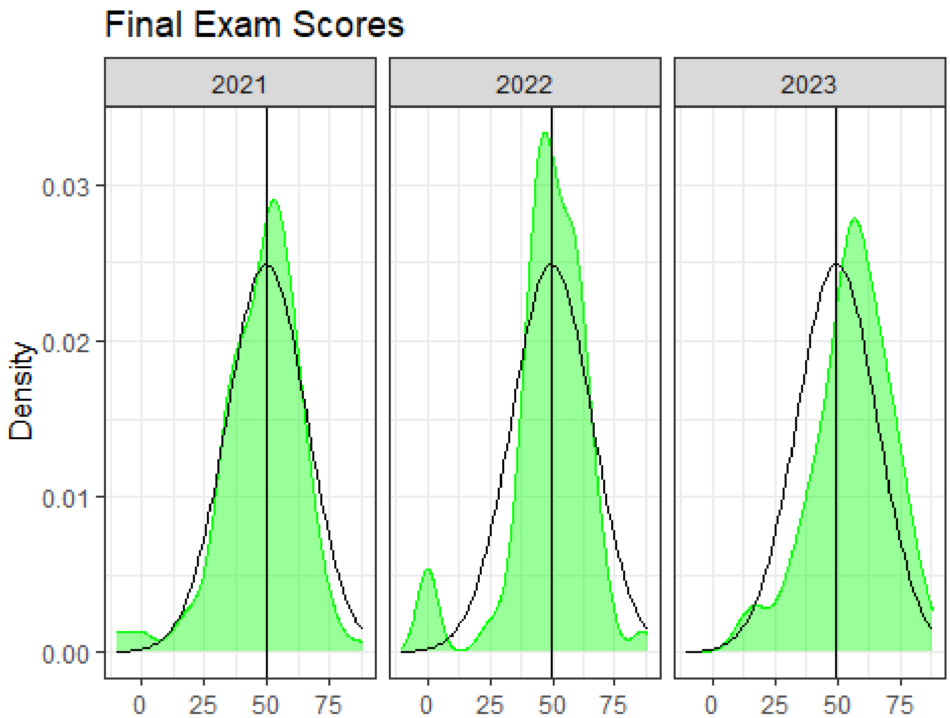

Figure 8: Final Exam score distributions across the all three conditions [2021: traditional lecture; 2022: traditional lecture with in-class active learning; 2023: traditional lecture, in-class active learning, and weekly 2-hour ALWS] plotted on x-axis. Density is shown on y-axis. Normal distribution curve is shown in black.

The Gaussian curve superimposed on the Final Exam score distributions here allows us to identify deviations more readily for each condition. 2021's distribution became more reflective of leptokurtosis, where the peak observed is greater than that of the normal distribution. This holds true for each of the observed conditions, with 2022 having the most dramatic example of this phenomenon, yet it also contains a second non-symmetrical peak, as was observed with the same condition during the Exam 2 analysis. The IQR of the 2021 condition drops considerably between Exam 3 and the Final Exam, from 31 points to 25. This indicates that the mean and median are more representative of actual scores. Considering the empirical values provided by

Tables 1 and 2, mean scores were: 61.55, 70.01, and 87.41 points for each condition; medians were 62, 71, and 91 points, for ’21 ’22 and ’23, respectively. Table 3 provides the p-value between the 2022 and 2023 condition as being p<0.001, which indicates that there is a statistically significant difference between the mean scores of the Final Exam among these two selected conditions.

Characteristic	N	**2021**, N = 144	**2022**, N = 70	**2023**, N = 66	p-value
Exam_1	280	58.50 (49.75, 69.00)	61.00 (51.25, 69.00)	60.00 (52.00, 68.00)	0.7
Exam_2	280	55.00 (45.00, 64.00)	56.00 (49.00, 64.00)	57.00 (46.00, 61.75)	0.8
Exam_3	280	55.00 (41.75, 68.25)	60.00 (50.50, 66.00)	58.00 (48.50, 66.00)	0.4
Final	280	50.50 (38.92, 58.08)	50.00 (44.00, 58.75)	57.00 (47.25, 66.50)	<0.001

Table 4 **:** Average failing scores of all four exams is show across all three conditions[2021: traditional lecture; 2022: traditional lecture with in-class active learning; 2023: traditional lecture, in-class active learning, and weekly 2-hour ALWS]. Q1 and Q3 are stated in parentheses. Scores across conditions in the same exam were compared using the Kruskal-Wallis Test. P-values are listed in the final column.

Table 4 demonstrates the median scores and IQR for the scores in the range of D’s, F’s and Withdrawals, also known as DFW scores. The Kruskal-Wallis Test was used to assess the significance of the changes between the three conditions, with the p-value reflecting a significant change by the Final Exam. This indicates that the 2023 condition’s median failing Final Exam score of 57 points was significantly higher than either 2021 or 2022’s and that the change is due to something other than chance.

Characteristic	**N**	**2021**, N = 144	**2022**, N = 70	**2023**, N = 66	**p-value**
Exam_1	280	58.06 (16.34)	59.46 (13.24)	59.53 (11.81)	0.7
Exam_2	280	53.81 (14.02)	54.74 (13.95)	55.02 (11.68)	0.8
Exam_3	280	54.31 (19.08)	58.09 (14.69)	56.35 (14.09)	0.4
Final	280	47.64 (15.77)	48.54 (15.91)	55.80 (15.40)	<0.001

Table 5: Average failing scores of all four exams is show across all three conditions[2021: traditional lecture; 2022: traditional lecture with in-class active learning; 2023: traditional lecture, in-class active learning, and weekly 2-hour ALWS]. Standard deviations are stated in parentheses. Scores across conditions in the same exam were compared using the Kruskal-Wallis Test. P-values are listed in the final column.

Similar to Table 4, Table 5 also displays the DFW scores, with emphasis on the mean and standard deviation of these values. It is important to note that the sample size changes across conditions as less students failed the class in 2022 and 2023 than did in 2021.

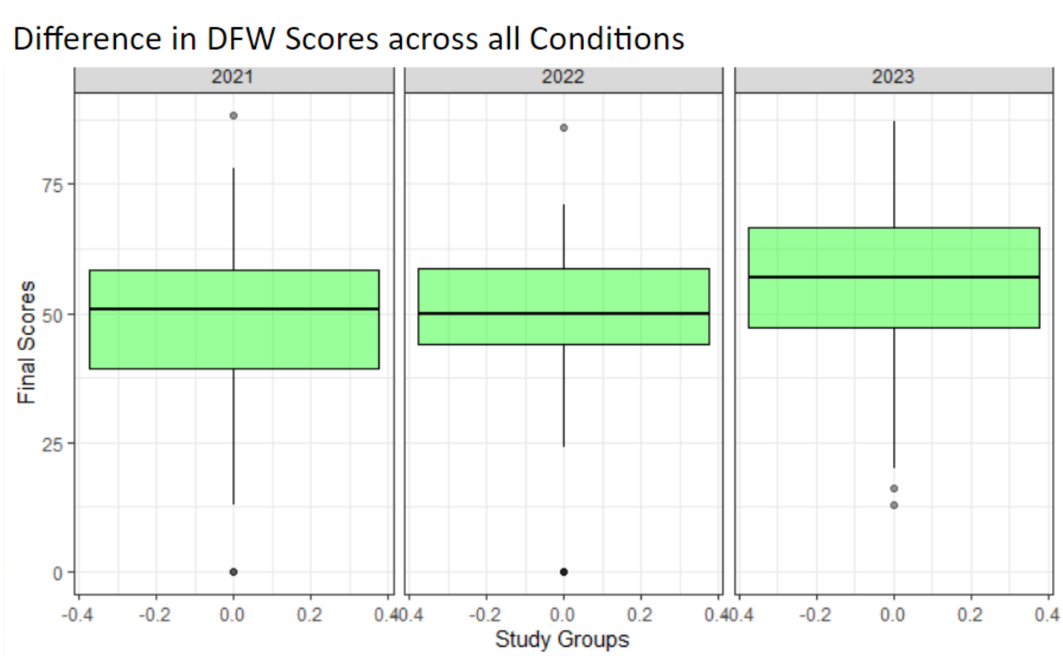

Figure 9: Final DFW score distributions across the all three conditions [2021: traditional lecture; 2022: traditional lecture with in-class active learning; 2023: traditional lecture, in-class active learning, and weekly 2-hour ALWS] plotted in a box-and-whiskers fashion. The final scores are plotted on the y-axis. Outliers are marked.

Figure 9 demonstrates the findings of Tables 4 and 5 visually and allows us to see how DFW rates changed over our observed conditions. The cohort of 2023 saw the greatest average failing score and the smallest sample size, meaning that less people failed the course overall.

Discussion:

We compared the means, standard deviations, medians and IQR's across all three years to study the impact of our conditions on outcomes. We found that even some active learning could create significant change, and that greater incorporations of active learning practices had on student performance even greater impacts on student performance.

We also analyzed separately between 2022 and 2023 to further refine the scope of the study. As discussed in the introduction, COVID 19 had a profound impact on the way students learn. We wanted to isolate and consolidate the data sets that were less extraneously impacted as the focus of this project is on active learning rather than mitigating the effects of asynchronous learning models. We also sought to understand whether the volume of active learning components affected student performance. By comparing two years where different styles of active learning were used, we were able to address this question. As we can see from Table 3, the p-value decreases with each subsequent exam suggesting that the improvement we are observing is less and less likely to be due to chance. This also supports the idea that ALWS work best as a consistent practice rather than an isolated intervention. Being that it is the first year of these course improvements, we must be vigilant in observing the changes and hopefully maintaining the trends we have observed above. In the coming years, it will be pivotal to track the progress of CH331 as it continues to evolve.

This project has shed new light on the importance of active learning practices, especially peer-led in STEM fields. As with any thoughtful inquiry, it generates many new questions in its wake, including how would interspersing active learning sessions more regularly into the week affect exam performance, and which other active learning practices might bolster student learning. The evolution of this project has taught us many things, highlighting the importance of teamwork on all levels of academia. Learning is not done in a vacuum, it is a circuit of presentation and application, affected by everything that makes us human. Acknowledging the emotional component of learning and working to humanize our approach to a seemingly intimidating subject may prove to be the key in unlocking any door. We must be willing to continuously evolve our teaching techniques with the ever-changing landscape of learning. New technologies and considerations are not a threat to our academic institutions, but rather a source of potential. This potential can only be tapped into when we refute tradition for tradition's own sake and seek to find new ways to tackle decades-old complexities. Sometimes, as in the case with this project, we come to appreciate the beauty in the simplest solutions. It is our responsibility to carve out time to consider new paths, as well as the emotional space to acknowledge that things can be done differently. Striving for continuous improvement in all aspects of academia is our duty to ourselves, our students, and generations of learners to come.

Bibliography

Anderman, E. M., & Murdock, T. B. (2007). The psychology of academic cheating. In E. M. Anderman & T. B. Murdock (Eds.), *Psychology of academic cheating* (pp. 1–5). Elsevier Academic Press. https://doi.org/10.1016/B978-012372541-7/50002-4

Bloom's Taxonomy - Center for Instructional Technology and Training - University of Florida. (2024). Retrieved June 7, 2024, from Ufl.edu website: https://citt.ufl.edu/resources/the-learning-process/designing-the-learning-experience/blooms-taxonomy/

Carter, A. J., Croft, A., Lukas, D., & Sandstrom, G. M. (2018). Women's visibility in academic seminars: Women ask fewer questions than men. *PLOS ONE, 13*(9). https://doi.org/10.1371/journal.pone.0202743

Code.org. (2016, June 16). What classes do students like the most?. ANYBODY CAN LEARN. https://blog.code.org/post/146020540698/what-classes-do-students-like-the-most

Grove, N. P., Hershberger, J. W., & Stacey Lowery Bretz. (2008). Impact of a spiral organic curriculum on student attrition and learning. *Chemistry Education. Research and Practice, 9*(2), 157–162. https://doi.org/10.1039/b806232n

Herr, O. (2024, February 15). Learning a language gives the study of STEM a little "je ne sais quoi." *British Science Association*. https://www.britishscienceassociation.org/blog/learning-a-language-gives-the-study-of-stem-a-little-je-ne-sais-quoi

Increasing Retention of Under-Represented Students in STEM through Affective and Cognitive Interventions. (2018). Retrieved June 7, 2024, from ACS Symposium Series website: https://pubs.acs.org/doi/pdf/10.1021/bk-2018-1301.ch006

Lee, N. (2018, January 18). The simple secret to groundbreaking science: work together. *Fourth Industrial Revolution*. https://www.weforum.org/agenda/2018/01/the-secret-of-groundbreaking-science-global-collaboration/

Libraryvoices. (2018, May 22). Library Voices. https://sites.uw.edu/libraryvoices/2018/05/22/approaching-organic-chemistry-from-someone-who-hated-organic-chemistry/comment-page-1/

Marnik, E. (2021, April 8). Making science more understandable. *ecrLife*. https://ecrlife.org/making-science-more-understandable/

Organic Chemistry, Life, the Universe and Everything (OCLUE): A Transformed Organic Chemistry Curriculum. (2019). Retrieved June 7, 2024, from *Journal of Chemical Education* website: https://pubs.acs.org/doi/10.1021/acs.jchemed.9b00401

SACNAS. (2020). Humanizing STEM: The Art of Storytelling. *YouTube*. Retrieved May 27, 2024, from https://www.youtube.com/watch?v=L5VE0Lbu0C8

Sridharan, S. (2018, February 26). You don't need to be a "humanities person" to study the humanities. *The Aggie*. https://theaggie.org/2018/02/25/dont-need-humanities-person-study-humanities/

Urban, S. (2017). Pen-enabled, real-time student engagement for teaching in STEM subjects. *Journal of Chemical Education, 94*(8), 1051–1059. https://doi.org/10.1021/acs.jchemed.7b00127

Yanai, I., & Lercher, M. (2020). The two languages of science. *Genome Biology, 21*(1). https://doi.org/10.1186/s13059-020-02057-5

Yanai, I., & Lercher, M. J. (2023). Make science disruptive again. *Nature Biotechnology, 41*(4), 450–451. https://doi.org/10.1038/s41587-023-01736-5