HOMOGENEOUS AND HETEROGENEOUS PEER GROUPS

IN THE SECONDARY SCIENCE CLASSROOM

by

Nathanael J. Dille

A professional paper submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Science Education

MONTANA STATE UNIVERSITY
Bozeman, Montana

July 2024

ACKNOWLEDGEMENTS

While designing, implementing, and analyzing my project, I have had a supportive team of fellow educators of various backgrounds to aid me.

I would like to thank Walt Woolbaugh, Ed.D, and Amanda Mello, who have served as advisors on my research project from the beginning, assisting me in refining my research questions, in designing my research instruments, and in effectively analyzing my data.

I am grateful to have had Dr. Nicholas Childs as my science reader. As one of my physics professors at Montana State University, he has periodically looked over my work and given advice on how to best analyze and communicate my findings.

I would also like to thank Christine Dille, my mother, who has been teaching high school at a private institution for over ten years, teaching both physics and writing. She is both knowledgeable about teaching and understands the context of a rigorous physics classroom, and has served as a sounding board for my research from the beginning.

Lastly, I would like to thank Colleen Dille, my wife, for lending her expertise in Child Developmental Psychology and her assistance in brainstorming implementation methods for my treatments, as well as the social and emotional implications of my treatment on students' learning experience.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

ABSTRACT

Group work is a regular part of lab-based science classes. While the benefits of learning with and from peers are well-documented, little research has shown how the composition of the groups shapes the learning experiences and outcomes of individuals. This study comparatively examines how grouping students homogeneously or heterogeneously by academic performance affects overall learning outcomes through assessment scores as well as student perceptions of learning and group cohesion. To investigate this, students from two AP Physics classrooms participated in two cycles of research, defined as an instructional unit each. The classes were divided into homogeneous groups of students with similar academic proficiency for one cycle and heterogeneous groups representing students of a range of proficiency for another. Both classes experienced both grouping styles, allowing their assessment scores and survey responses to be compared between cycles. The results show that assessment scores were mostly unaffected by the students' group composition. On the other hand, student perceptions, as measured through student surveys given at the conclusion of each cycle, revealed that students report higher engagement and collaboration with peers while working in homogenous groups, thus implying that students feel more involved and interested in learning the science content when discussing and working primarily with peers of similar academic understanding.

CHAPTER ONE

INTRODUCTION AND BACKGROUND

Context of the Study

The goal of this action research is to address how secondary science students respond to different types of grouping. In the science classroom, students regularly work with peers to complete assignments, engage in discussion, or collect data during labs. Group work plays an important role in science curriculum, and therefore creating groups that better foster learning and engagement in science content is paramount to the success of subsequent group activities. For this reason, the following study will seek to provide practical strategies for how to group students.

In my classes, I position student desks in groups of four, regularly utilizing these groupings for quick discussions, longer lab activities, and even extended projects. In order to challenge students to work with diverse peers, I change the groups after every unit of study. While I have tried different strategies for grouping students, I often find that in a given group, one or two of the students completes the majority of the work or dominates the discussion. The others, often students who are less vocal or more challenged by the current content, take a back seat, comfortable with letting a more capable peer get the answers right on their behalf, resulting in the students who are struggling most learning the least during a group assignment.

Creating groups that engage every learner in science content is a necessary step to creating an equitable learning environment that challenges every student to grow. Not only will this study better prepare me for organizing my own students, its implications could extend to other high school classrooms. Group discussions and projects are not unique to science, and any strategy that increases student engagement, especially for students who may be struggling, is a valuable tool. While the research conducted will apply most directly to secondary science students, it will also contribute to the general literature on collaborative learning and group organization.

Focus Statement/Question

Traditionally, many educational writers suggest using heterogeneous groups—groups that include students of all different proficiency levels—to foster and encourage peer learning and ensure that all of the groups can move at a similar pace (Vygotsky, 1978). While I have used this strategy before, I have also noted some weaknesses in practice, as outlined above. As an alternative, students can be grouped homogeneously by placing the highest achieving students together and the lowest achieving – and often least motivated – students together. While this has the benefit of making differentiation easier and challenging the less motivated students to rise to the occasion, little research has directly compared these two styles in a science classroom. For this reason, my research question was, What are the effects of grouping students homogeneously or heterogeneously by academic performance on classroom engagement and learning outcomes in a secondary science classroom?

I have often noticed in my classroom that all students seem to be challenged in different ways when grouped with peers of similar academic proficiency. Groups of high-achieving students tend to compete with one another, sometimes even racing to complete a problem first. Groups of less motivated or less confident students require more individuals to contribute in order for a learning task to be completed, as there may not be one student in the group who can or will complete it by themself.

Students also pay attention to the groups they are placed in, and some students may attempt to figure out why they were placed in that group. When homogeneous groups are formed, students may feel more comfortable in their group because they are more likely to be surrounded by like-minded peers with similar skills and challenges. On the contrary, students may feel like they have been "typed," and sorted into a group that they do not want to be part of. For example, a student may realize that there are high-achieving groups in the room, and be upset that they are not placed in one of those groups. These areas of interest have led me to two sub-questions which will supplement my findings.

My sub-questions include the following:

1. What is the impact of homogeneous grouping on different subgroups and individual students? For example, are high-achieving students affected differently by homogeneous grouping than low-achieving students?
2. How do the different grouping styles influence students' perceptions of peer learning and group cohesion, both overall and in each subgroup?

CHAPTER TWO

CONCEPTUAL FRAMEWORK

History of Group Learning

Over the last few decades, group work and collaboration have become foundational aspects of learning, both in theory and in practice. Lev Vygotsky, a pioneer in developmental psychology, argued in 1978 that learning takes place in the context of relationships and interactions between individuals (p. 57). In the science classroom, working in groups, especially for labs, has become the routine for many teachers. Collaborative learning, however, begs several logistical questions: how should students be divided into groups, and are some methods of grouping students more conducive to learning than others?

In 1994, Johnson and Johnson began to answer these questions by postulating that collaborative learning is successful when "each group member's efforts are required and indispensable for group success" (p. 2). Since then, several studies have found a correlation between the quality and frequency of group interactions and individual members' performance on subsequent assessments, suggesting that groups learn most successfully when they work together (Cohen et al., 2002; Schultz, 1999). While personality differences and unpredictable factors can certainly affect the cohesiveness of a group, steps can still be taken to construct student groups that foster collaboration. Rachel Lotan (2003), one of the researchers from the 2002 study, states that "positive interdependence among members of a group" is one of the predictors of learning within the context of a group (p. 74). For this reason, one of the goals of this study was to examine what factors contribute to positive interdependence.

Positive interdependence within a group can only be realized if all members of the group participate in the given activity. One notable downside to allowing students to work together is the difficulty in ensuring that each member contributes to the discussion or final product. While there are numerous strategies and designs that can encourage participation from all parties, a recent study on the efficacy of group educational activities found that a student's engagement in the activity was a powerful mediator between the activity's design and the student's perceived learning outcomes (De Hei et al., 2018, p. 2363). As student engagement serves as a predictor of learning outcomes, efforts must be taken to construct learning groups that maximize the engagement of every member of the group.

Grouping by Academic Performance

Traditional wisdom in collaborative learning environments has been to group students heterogeneously by academic performance. This mentality stems from Vygotsky's theories. Vygotsky (1978) believed that learners acquire new problem-solving skills "under adult guidance, or in collaboration with more capable peers" (p. 86). While the impact of the teacher on learning is indisputable, many teachers have taken the second part of that to mean students should regularly be grouped with students of differing learning levels, in order to encourage peer learning (Lewis, 2019; Manfre, 2022). Unfortunately, there are a few challenges with taking this approach.

Heterogeneous groups face three distinct challenges in the science classroom. First, these mixed-ability groups assume high-performing students are willing and able to pass on their knowledge to their peers. While this is sometimes true, not all gifted students have the desire or skillset to teach others (Clinkenbeard, 2012, pp. 627-628). The second challenge for heterogeneous groups is only the lower-performing students have the ability to learn from their peers, while the high-performing students have no more capable peers from which to learn. The third, and perhaps greatest problem, is the effect of student hierarchy within the group. Research into group dynamics has revealed that even when individuals are randomly assigned groups, they form a "status ordering," where a student who is "relatively quiet is seen as having made the least important contribution and is felt to have contributed few good ideas" (Cohen & Lotan, 2014, p. 28). This can be seen in a classroom lab activity where a high-performing, louder student completes the majority of the lab, while a quieter, less motivated student watches and learns very little from the activity.

In order to combat these problems, students can be organized into homogeneous groups by academic performance. This means high-achieving students are grouped with fellow high-achieving students, and low-achieving students are likewise grouped with students of a similar performance level. This approach to grouping would combat the first and third problems mentioned above, as the higher-achieving students may not experience the pressure to pass on their knowledge, while the lower-achieving groups will not be dominated by one knowledgeable student. In 1999, both styles of groups were tested in math classes, and their results were compared for a small group of high-achieving students. The researchers discovered that the students' summative performance "was greater for all students in the homogeneous setting compared to their heterogeneously grouped peers" (Kanevsky & Sheppard, 1999, p. 270). Unfortunately, this study did not examine the effects on low-achieving students, nor does it take place in a science classroom, which can rely more heavily on groups for labs and activities than math classes.

One notable challenge of homogeneous groups, is that they rely heavily on the teacher to support and challenge students, particularly in the low-achieving groups, acting as the "more capable peer" (Vygotsky, 1978, p. 86). A recent study gathered data from mathematics teachers using either heterogeneous or homogeneous groups. The study found that all teachers thought the type of grouping they currently used promoted higher engagement and learning outcomes, stressing "the role that the teachers have on the efficacy of grouping students according to the level of ability" (Cotic et al., 2023, p. 2). Not surprisingly, teachers who are experienced in one of these styles of teaching are able to support students in the style of groups with which they are familiar. For that reason, the focus of this action research project was to test the performance and social outcomes of using heterogeneous and homogeneous groups for the same students in a science classroom and under the same teacher.

CHAPTER THREE

METHODOLOGY

Demographics

All data collected for this research took place in an AP Physics 1 classroom. AP Physics 1 is a first-year physics class created by the College Board and offered at Naperville Central High School (NCHS). NCHS is a suburban high school in Naperville, Illinois, about 45 minutes outside of the city limits of Chicago. Currently serving a student population of 2,612 students in grades 9-12, NCHS is very highly ranked, both by a graduation rate of 97% and through state-wide testing (Naperville Central, 2021). Naperville as a whole is considered a wealthy suburb. The median household income in Naperville is $127,648, and the U.S. Census Bureau only categorizes 3.8% of the population as persons in poverty (Naperville City, 2021). Within NCHS, the population is predominantly White and Asian, with population percentages of 66% and 17%, respectively (Naperville Central, 2021).

This study was completed in two AP Physics 1 classes. As an elective course, AP Physics 1 tends to serve students who are interested in science or are high-achieving students who regularly take honors and AP classes. Students who are less interested in physics have the option to take general physics or other science classes, such as earth science or astronomy. Nevertheless, AP Physics 1 remains a popular class at the school, as currently 32% of the 11th grade class is enrolled.

In the first class, which will be referred to as class A, there were 23 students, 11 of whom identified as female and 12 who identified as male. Of these students, 17 consented to participate in the study. In the second class, which will be designated as class B, there were 22 students, six of whom identified as female and 16 who identified as male. Eighteen students from class B participated in the study. While the course was primarily designed for Juniors (11^{th} grade) who are taking physics for the first time, there were two students in 10^{th} grade in class A, and two in 12^{th}. Similarly in class B, there were four students in 10^{th} grade, and one in 12^{th}. There was one student in each class with an Individualized Education Plan (IEP) and/or a Section 504 plan. There were four students in class A and one student in class B who recently received English language support courses. All students in the class were either concurrently enrolled in pre-calculus or had already completed it.

Class A met during first hour, at 7:45 am each morning, for 45-50 minutes, depending on the day of the week. This means the students in the class started their day with physics, and despite the early hour, the classroom culture tended to be fairly energetic. Class B met during the eighth hour. Physics was their last class of the day, and they were always very lively. I often started classes with a short question, usually either a physics warm-up problem or a community building, ice-breaker question. After the opening routine, I tried to vary the class time between lab explorations, problem solving activities, and lecture-based note-taking.

Treatment

To form homogeneous and heterogeneous groups, students were grouped based on previous academic performance. Specifically, each class was be divided into (roughly) six groups of four students. In my treatment class, the eight students with the highest current grades were placed into one of the two high-achieving groups, the eight students with the lowest current grades were be placed into low-achieving groups, and the remaining eight students formed the mid-achieving groups. Students were not informed as to the method of grouping or the reason behind their placement. For the non-treatment class, I followed the same process, but I intentionally formed groups that represented a least one high, medium, and low-achieving student.

In order to compare the two different styles of grouping, the two classes were observed simultaneously during two distinct research cycles, representing two units in the physics curriculum. For most comparable results, the first cycle took place during a unit on introducing forces and Newton's laws of motion, and the second unit was the application of forces and motion to new and unique scenarios. For the first cycle, class A was considered the treatment class and organized into homogeneous groups by academic performance, while class B, the non-treatment class, was organized into heterogeneous groups by academic performance. After one unit of study was concluded and data was collected, class A became the non-treatment class and class B the treatment class for the second cycle. Each cycle took approximately three to four weeks, and incorporated two summative labs, three quizzes, and one test. This process of switching the treatment and non-treatment groups is illustrated below in Figure 1. Completing these two cycles should minimize the effects of differing ability levels within the classes on the data, as all students were observed in both the treatment and non-treatment.

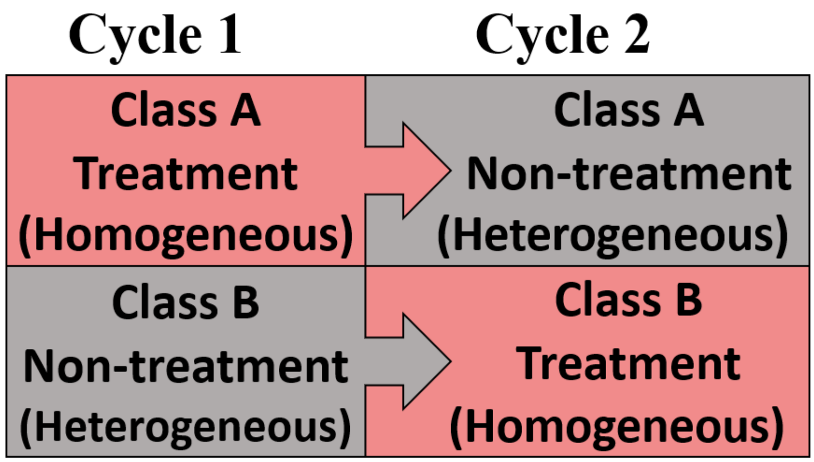

Figure 1. Treatment and non-treatment class during each research cycle, illustrating how each class spent one cycle as the treatment group and one cycle as the non-treatment group.

During each cycle of study, students engaged in a range of assessments. Group activities included regular discussions and problem solving, several quizzes, as well as two interactive laboratory experiments, completed as a team. Student scores on each lab were recorded individually by a follow-up lab task or quiz, and students were also individually scored at the end of the unit on a summative test. Individual scores were compared and used to determine the effects of the different grouping styles on learning at the conclusion of each cycle.

As an additional measure of student engagement, students completed two surveys at the end of each cycle. The surveys included Likert-scale peer-review and self-review questions, prompting students to reflect on their learning and collaboration. Students were informed that the post-cycle survey would not contribute to their grade or their groupmates' grades in any way, as previous research in peer assessment has revealed that students tend to score their peers "without bias when the mark does not count towards the final grade" (Boud et al., 2019, p. 864). In order to collect accurate and unbiased data on the cohesiveness and efficacy of groups, students were made aware that survey responses would be anonymous and only used for the purpose of the study.

The post-cycle surveys used were inspired by Ryan Foley's (2012) study of differentiated groups and their impact on student engagement (p. 9), and are included as appendices A and B. Using Likert-score questions, the contribution of each student in the group was quantified by averaging the selections of all students in a group. These surveys also served to corroborate data collected through assessments by collecting qualitative student reflections, providing an opportunity to compare the effects of different grouping styles on collaboration, teamwork, and attitudes toward learning in the different types of groups.

Data Collection and Analysis Strategies

Multiple types of data were collected to answer each research question. Because engagement can be an elusive concept to measure, student engagement was assessed primarily by assessment scores. In order to accomplish this, a variety of assessments were used, all of which were an ordinary part of the AP Physics 1 curriculum. The most straightforward data collected was the quiz and test scores from each unit. Each research cycle coincided with a unit of study in class, which means there were two quizzes, serving as checkpoints in the unit, as well as a test at the end. In addition to these assessments, students completed two lab investigations in each unit. As a rule, the AP Physics 1 teachers, including myself, always gave an individual assessment at the end of each lab to ensure that lab grades reflect the individual student's understanding, and not that of the group as a whole. These lab assessments vary, but may take the form of a lab quiz, a report, or a worksheet, all of which received a numeric lab score, and can be compared between treatment and non-treatment groups, or between various subgroups.

Comparing student scores on the assessments, however, presented a population and content challenge. If average student scores in the treatment class were directly compared to the non-treatment class during a given cycle, a difference in the classroom culture or population of students could constitute a bias in the results. If only one class of students' scores were compared to their own scores from a previous or subsequent cycle, the population bias could be eliminated, but the inability to ensure that the content of each unit is of the exact same difficulty introduced a confounding variable. To combat this dilemma and ensure the reliability of my findings, the treatment scores and responses from each cycle were compiled, and the non-treatment scores were compiled. Because each class served as the treatment group for one cycle and the non-treatment group for the other, this ensured that each data set represents both classes' participants, as well as content from both cycles, minimizing the effects of population or content differences.

While the assessments mentioned formed the backbone of the data used to measure engagement, quantitative and qualitative data were also gathered from the aforementioned surveys (appendices A and B) to supplement my findings, triangulate my data, and answer the second sub-question, pertaining to students' perceptions of group learning. Both surveys were given as Google Forms to all students at the close of each cycle, meaning each student eventually took each survey twice. The Likert-style questions in the self-review survey were designed to assess whether a student felt more engaged in one unit than another and how often they worked with their group in each cycle. While the answers given in the first cycle were interesting to review, the most useful analysis of the survey responses was comparing students' own responses from one cycle to the next, reflecting any changes in perceptions from treatment to non-treatment or vice versa.

The peer-review survey was intended to function as a reliability check on other data, to corroborate or critique the findings of the other instruments. Students were asked to report on how much they contributed to the group's work as opposed to the other members of their group when working on labs, as well as how many of their groupmates were regularly disengaged during class. These responses did not only serve as an indicator of the effects of my treatment/non-treatment on group collaboration, but also allowed me to spot any groups that were particularly incohesive, in case I wanted to identify any groups as outliers.

The open-ended question on each survey also provided qualitative data on student perceptions. By asking students to share their reasoning, this question allowed students to share their experiences, and perhaps fill in any gaps in the questions I chose to ask in the surveys. The various data-collection tools are summarized in Table 1, along with the type of data they will collect and to which research question they pertain. Note that each of the research questions is answered initially with quantitative data, with qualitative data used to supplement and corroborate findings as much as possible. By providing at least three separate types of data to answer each question, I was able to use triangulation to check that the differences in assessment scores between classes and cycles were substantiated both by scores on diverse types of assessments and by students' perception of learning.

Table 1. Summary of data collection methods for each research question.

Data Collection Matrix	**Quantitative data**	**Qualitative data**
research question:		
Research question: What are the effects of homogeneous vs. heterogeneous groups on engagement and learning outcomes?	1. scores on summative unit assessment 2. individual student scores on labs 3. self-review and peer-review survey ratings	1. self-review survey reflection questions
Sub-question 1: What is the impact of different grouping styles on subgroups and individual students?	1. scores on summative unit assessment 2. comparing scores on labs across subgroups	1. comparing self-review reflections across subgroups
Sub-question 2: What is the impact of different grouping on group cohesion and student perspectives?	1. peer-review survey ratings 2. self-review survey participation ratings	1. cross-referencing peer-review survey responses within each group

All instruments used in this study have been approved by the Institutional Review Board of Montana State University, and the approval is included as Appendix C, given that student assent and signed parent/guardian consent was obtained due to the use of confidential student information by the researcher. The assent/consent forms are included in Appendix D, and all subsequent data was obtained solely from participants who turned in assent/consent forms.

CHAPTER FOUR

DATA ANALYSIS

Assessment Results

The first stage of my data analysis compares the quantitative data gathered during the two cycles, looking for relationships between the treatment and the scores on tests, labs, and quizzes. Comparing scores in a given class between cycles, however, may not properly account for the prospective difference in difficulty between the two units of study. Likewise, comparing the treatment and non-treatment class in the same cycle would not account for the population differences within the classes. For these reasons, all scores from the treatment classes in both cycles were compiled to make the treatment data, while all of the non-treatment scores were compiled to form the non-treatment data. This way, each data set includes scores from class A and class B, and also from cycle 1 and cycle 2.

After compiling scores, the mean scores for each assessment type were calculated and compared, as shown in Table 2. Overall, these data show that the treatment had little to no effect on the overall scores on assessments. After running a t-test on each category, it was determined that the difference was not statistically significant in any type of assessment used in this study ($p>0.05$).

Table 2. Mean score on each assessment category, (*N*=35).

Assessment Type	Treatment	Non-treatment
Mean Test Score (%)	78	78
Mean Lab Score (%)	88	89
Mean Quiz Score (%)	88	87

While these initial results were not very exciting, the next step was to compare how the treatment affected specific subgroups within the population. Because students were already divided into high, medium, and low achieving categories in order to determine grouping, the treatment and non-treatment scores were compared for each of these subgroups. Figure 2 shows the comparative test scores of the treatment and non-treatment groups divided into subgroups. As before, none of the variance were determined to be statistically significant, but there are some noticeable differences at this level.

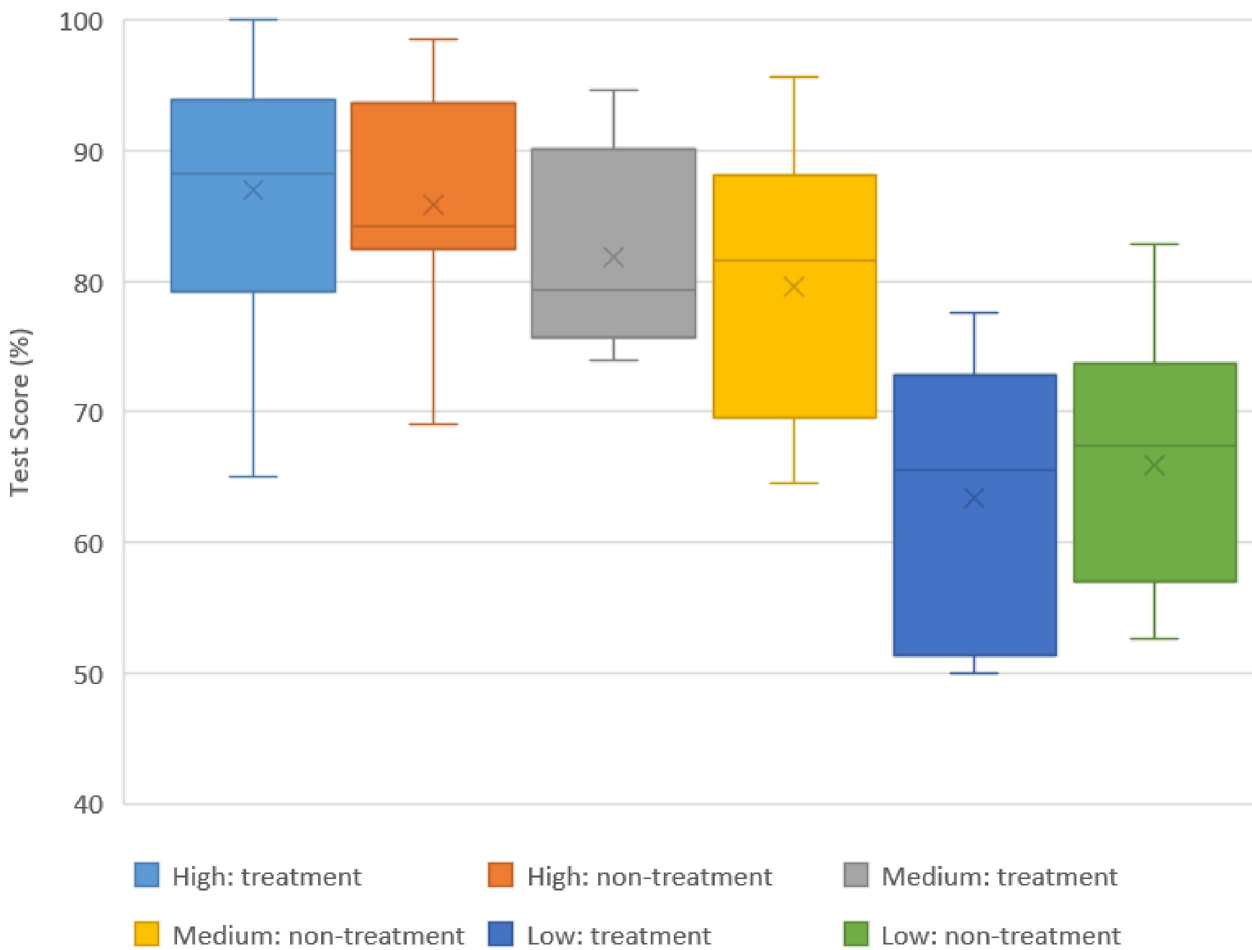

Figure 2. Comparative test scores for each subgroup, where high, medium, and low refer to the academic proficiency of the subgroup, (*N*=35).

As seen in the box-and-whiskers plot (Figure 2), the treatment had little effect on the test scores. For the mid-achieving students, the mean score did move down slightly on the non-treatment assessments, and the mean score moved up slightly on non-treatment assessments for low-achieving students, but not enough to state with confidence that this was due to the treatment. Results were similar when comparing scores within subgroups for the other assessment types, signifying that overall, homogeneous and heterogeneous groups have minimal effects on students' learning. To determine whether student engagement is affected by the treatment, these assessment results will be compared to student survey responses from each cycle.

Survey Findings

Another way student engagement was measured was through self-review and peer-review, collected by surveys (Appendices A and B). Each participant reflected on their engagement and that of their peers at the end of each cycle, thus providing responses for each class while in each cycle. After all treatment responses from both classes were combined, it was found that overall self-reported engagement seemed to be higher for each class while in homogeneous groups, or the treatment. Figure 3 shows the responses for each class while in the treatment (homogeneous) and non-treatment (heterogeneous) groups.

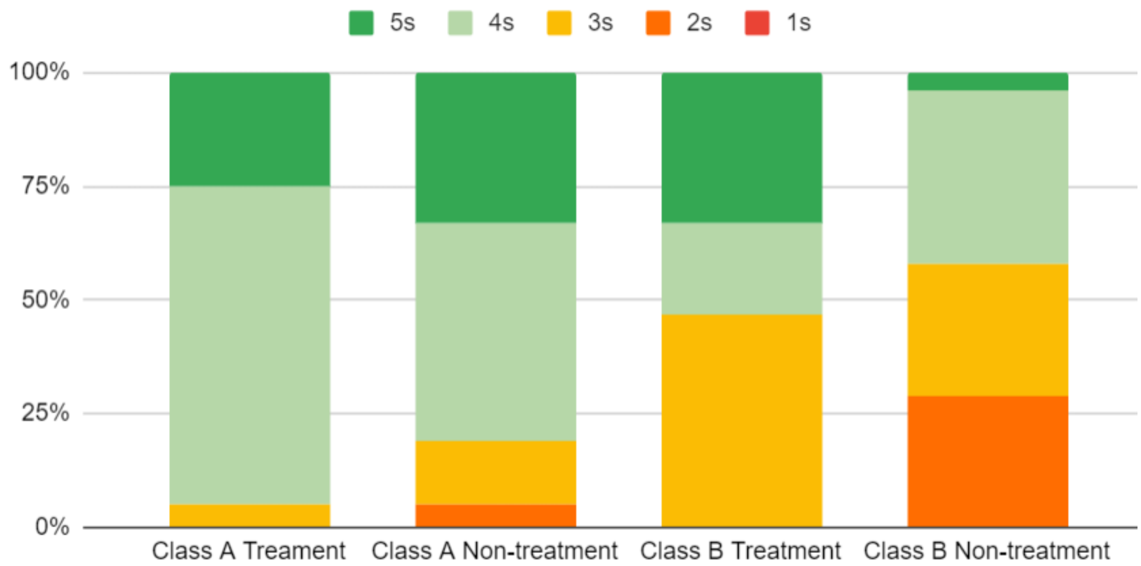

Figure 3. Self-reported engagement on a scale of 1-5, with 5 meaning almost always engaged during class and 1 meaning almost never engaged during class, (*N*=35). Overall trend suggests that each class was more engaged during the treatment cycle.

While class A seemed more engaged than class B in general, both classes self-reported a higher average engagement during the treatment cycle. More specifically, the number of students from class A (*n*=24) who rated themselves as 3 or lower rose from 5% to 19% from treatment to non-treatment, while in class B this statistic rose from 47% to 58%, with an increase of 29% in students reporting a 2. This definitely suggests a decrease in student engagement, though a chi squared test did not find the difference significant enough to make a claim with certainty ($p>0.05$). Because the classes alternated which class was the treatment and non-treatment, variables such as the content of each unit itself can be ignored, meaning the change in response can be primarily contributed to the different groups that students worked with.

The peer-review survey asked students to report how many members of their group (not including themselves) were regularly disengaged during class time. This statistic also appeared noticeably different from treatment to non-treatment. The results are shown in Figure 4, and suggest that the number of disengaged students was always higher while students were in heterogeneous groups. While the difference is notable, a chi squared test once again revealed that the difference was not sufficient enough to make a definitive claim.

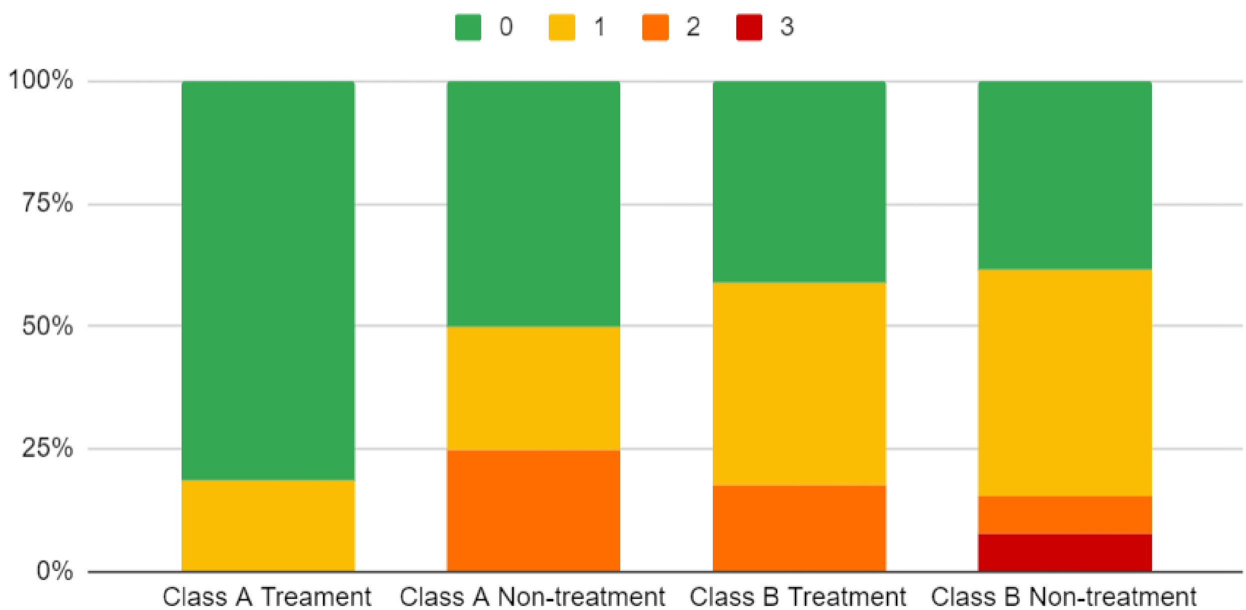

Figure 4. Student responses to “how many members of your lab/table group were regularly disengaged?” Note that the number of students who responded with 1 or more was higher in non-treatment cycles than in treatment cycles, (*N*=35).

Once again, to better understand students’ responses, the data were divided into subgroups, allowing the comparison of how different populations of students reported on their engagement and the engagement of their peers. While the overall trends often followed the same patterns, there were also some notable differences in the subgroups. By chi squared analysis, the difference for any subgroup was not statistically significant enough to reject the null hypothesis ($p>0.05$), but the number of students who reported a 3 or lower increased from treatment to non-treatment in all subgroups individually, with mid-achieving learners showing the most dramatic difference.

When looking at number of peers who are regularly disengaged, it was also noted that all subgroups reported more disengaged peers while in non-treatment than in treatment groups. For this response, the low-achieving demographic gave an even more dramatic difference. Though still not statistically significant ($p<0.05$), it suggests that there was a notable difference in the number of disengaged students reported by this subgroup.

In the self-review survey, participants were also asked to report on a scale of 1-5 how often they worked with peers to solve or discuss a problem, with 5 meaning they almost always worked with a peer from their group. Notably, students consistently answered that they worked with peers more while in homogeneous (treatment) groups than in heterogeneous (non-treatment) groups, as illustrated in Figure 5. While the difference is less pronounced, and not enough to make a definitive claim, it still points to an overall trend of homogeneous groups correlating to more collaboration among peers.

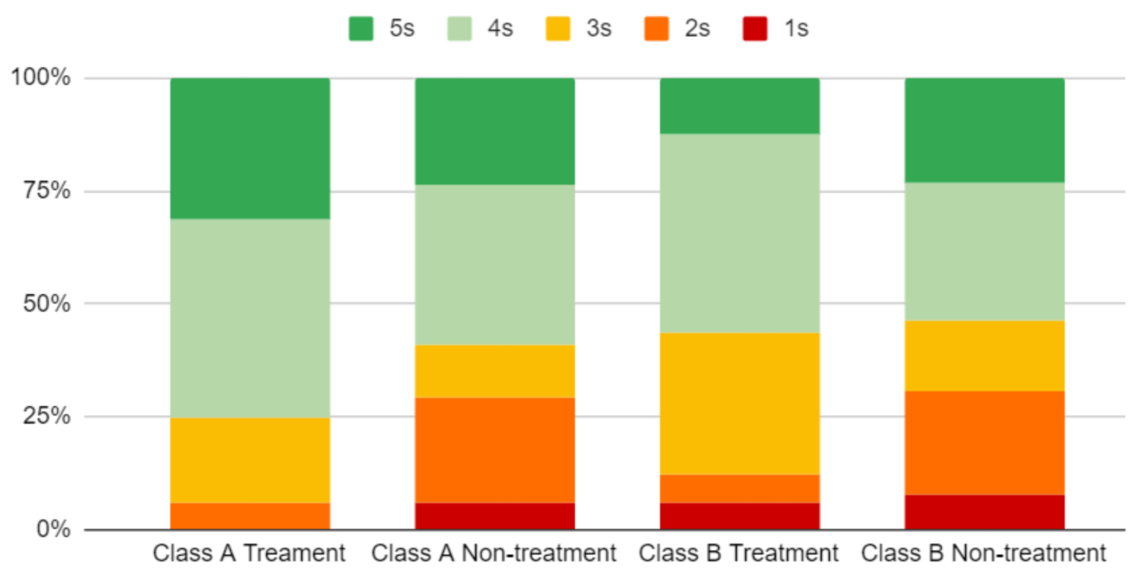

Figure 5. Self-reported collaboration on a scale of 1-5, with 5 meaning almost always working together with group/table members and 1 meaning almost never working with group/table members. Data suggest that students collaborate more while in homogeneous (treatment) groups, (N=35).

Looking more carefully at how different subgroups reported on collaboration, it was found that high-achieving learners reported the most significant drop in collaboration from treatment to non-treatment. Figure 6 gives the results for this specific subgroup. By chi squared analysis, the difference cannot be considered definitive, but the change in response from homogeneous to heterogeneous is still noticeable, and showing a greater willingness to collaborate with peers of similar academic performance.

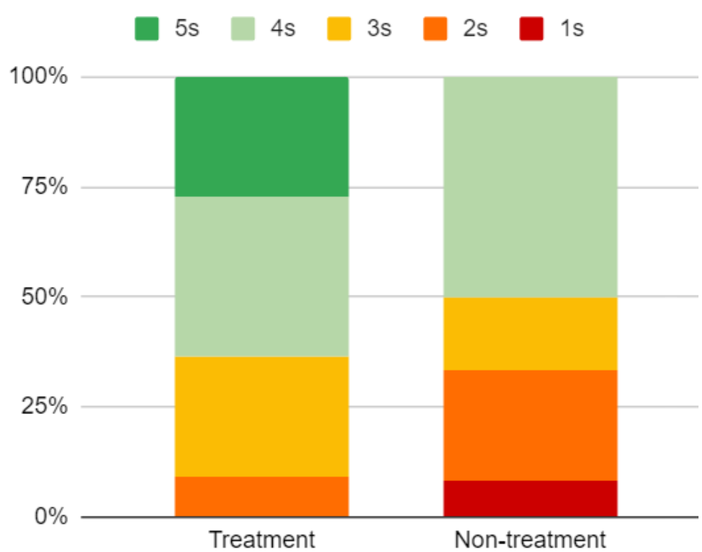

Figure 6. Self-reported collaboration from high-achieving students on a scale of 1-5, with 5 meaning almost always working together with group/table members. High-achieving students seem to collaborate more with other high-achieving students, (*n*=12).

To fully assess the effects of the treatment on cohesion and collaboration, students were also asked how much they personally contributed to their group during labs and assessments. They could answer that they contributed more, less, or about as much as the other members of their group. The data in Figure 7 show responses to this question, indicating that students felt like their contribution was more on par with their peers in homogeneous groups. Notably, the number of responses of students saying they did either noticeably more or less than their peers both rose for heterogeneous groups.

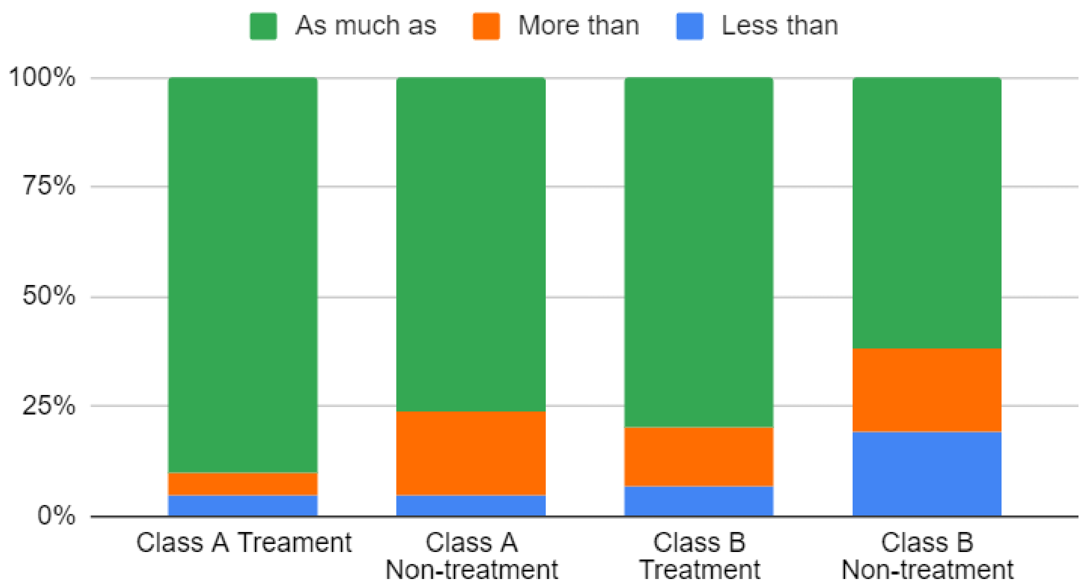

Figure 7. Student responses to “how much of the group work did you personally contribute?” Note that the number of students who reported that they did either noticeably more or less work than their peers was higher in non-treatment cycles than treatment cycles, (*N*=35).

Overall, while few of the data sets yielded a statistically significant difference in responses from treatment to non-treatment, the volume of data that all suggest a similar trend cannot be ignored. In all data sets, students reported that they feel more engaged, collaborate more, and take more of an active role in homogeneous groups than heterogeneous groups on average. Not only was this true for the total population that participated in the study, but the same trends were found in each subgroup individually, strongly suggesting a link between the treatment and student perceptions of engagement and collaboration.

CHAPTER FIVE

CLAIM, EVIDENCE, AND REASONING

Claims From the Study

The goal of this research was to test the effects of different grouping styles on student engagement and on learning. After analyzing my data, I believe there are two parts to this question that must be answered separately. Learning outcomes, as reflected through assessment scores, were minimally affected by the treatment, if at all (Table 2). This was true both for overall scores, and for all subgroups separately. On the other hand, the data do suggest that the treatment has a positive effect on student engagement, both overall, and in each subgroup. For all subgroups, students reported that they were more engaged and that less of their peers were disengaged when they were grouped with peers of similar academic performance (Figures 3 & 4). In these homogeneous groups, students also seemed to relate to and collaborate with their peers more readily. While statistically, the difference was still small, the number of responses to various survey questions that supported this overall trend holds some weight. Not only is this evidenced by the quantitative survey data, but student responses to the open-ended survey questions corroborated the results and uncovered several themes, some of which were unique to specific subgroups.

For high-achieving students, many of the responses mentioned how their group worked together better in homogeneous groups and that they didn't feel challenged by their peers when in heterogeneous groups. For example, one student from class A said she was less likely to "zone out" and that "my group was very collaborative" when in the homogeneous (treatment) cycle, while a student from class B expressed frustration that "some people in my group did not care or were not engaged in trying to learn" while in the heterogeneous (non-treatment) group. This pattern seemed to equally weigh on low-achieving students in a different sense. One of these students said he found it hard to keep up in the heterogeneous group, so he would "wait till someone gets the answer." One of the lower-achieving students from class A said, "whenever I had a question, the people in my group wouldn't help me that much," suggesting that the high-achieving students did not always have the willingness or required skillset to assist their peers, which was substantiated by Pamela Clinkenbeard in 2012 (pp. 627-628). While there can certainly be cases where students adopt a peer-tutoring role in heterogeneous groups, and some students responded positively to these groups (one student said her heterogeneous group worked together "efficiently and effectively"), more often it seemed that all students in homogeneous groups were less willing to collaborate.

This lack of cohesion and collaboration when working with peers who have a significant difference in content understanding was found to affect all students in the group. In response to my research question on student perceptions of peer learning, I found that both high and low-achieving students often expressed frustration with each other in heterogeneous groups, which leads to less productive collaboration for all members, including mid-achieving students. In homogeneous groups, on the other hand, students may have felt more confident that their contributions were valuable and necessary. Knowing that their peers are all at a similar level of understanding may have encouraged positive interdependence between all members (Lotan, 2003, p. 74). This is evidenced in Figure 7, where the number of students who answered that they contributed about as much as their peers, indicating an equitable, interdependent group dynamic, was always higher in homogeneous (treatment) groups.

Despite the overall positive response toward homogeneous groups on student surveys, there were also some students who disagreed, particularly in one low-achieving group. These students felt that none of them had enough conceptual understanding to get started on a task. One of the members mentioned that his group rarely worked together when not required because they "aren't smart enough," perhaps indicating that he was catching on to the treatment. While this is a response no teacher wants to hear, it points out a possible negative reaction to homogeneous groups, and one that must be combatted by the teacher's role in group activities.

It is worth noting that the role of the teacher in the different grouping styles is substantially different, especially during group work, which can influence the efficacy of each method. During heterogeneous (non-treatment) classes, the teacher's role is often to stimulate discussion and challenge students, spending time roughly equally around the room with each group. During homogeneous (treatment) classes, the teacher can choose to differentiate the lesson by spending more time working with low-achieving groups, offering them additional support and ideas, and spending the least time with the high-achieving groups, challenging them to work more independently and generate more solutions as a team. This built-in feature of homogeneous groups can ensure that low-achieving students are still able to rise to the same challenge as their peers.

Value of the Study and Consideration for Future Research

One notable limitation of this study is that all participants were in AP Physics classrooms. This means the findings discussed above apply most directly to honors or AP students in high school science classrooms. While many other types of classrooms use student groups for activities, projects, and daily work, I believe the results could be very different if the study were conducted in a general level classroom. General level classes often represent a broader range of students, both in academic proficiency and in behavior, which often lead to very different peer relationships and interactions. Because homogeneous groups are designed to maximize the opportunity for instructional differentiation by allowing different students to work at different paces in the same classroom, the effects, and even benefits, of homogeneous grouping could be more pronounced in a more diverse classroom, but further research would be needed to corroborate that hypothesis.

Another consideration for future research would be the impact of placing students in homogeneous groups over a longer period of time, such as multiple school years. While most teachers at my high school either do not use student groups regularly or create groups based on different criteria, using homogeneous groups prolifically could create a unique social structure. While generally speaking, high school students seem to be self-aware of their own strengths and vices, younger students could be shaped by the social groups formed by proficiency-based grouping. It may be that these effects would not differ significantly from the social structures that form naturally, but the long-term psychological and social effects would need to be investigated before homogeneous groups can be recommended for younger age groups.

Impact of Action Research on the Author

As the teacher and researcher during this study, I noticed many benefits of homogeneous grouping, as well as cases where it was not as effective. While sometimes students were encouraged to work with peers of similar abilities, I also found times where high-achieving groups preferred to work individually or low-achieving groups did not feel up to the task. Finding a group dynamic that is efficient for learning and encourages positive, interactive peer relationships is a challenge, but also a goal for my classes as I continue to search for reliable ways to build a supportive, safe, and inquisitive community in my classrooms. The overall positive student perception of homogeneous groups despite their minimal effects on assessment scores was surprising, and I intend to explore the variables and limits of these grouping structures more within my classes in the future.

While there are significant benefits to be appreciated about designing intentional student grouping strategies that maximize the learning experience for all students, I also learned that the teacher's knowledge of their students is still paramount and irreplaceable. Some students have unique needs or work best with (or without) specific students. Some students just prefer working alone, regardless of their proficiency in the class or who is around them. In the end, student groups have a large impact on the classroom culture, but that impact is based on many factors within the group, with academic proficiency being one of them.

REFERENCES CITED

Boud, D., Sridharan, B., & Tai, J. (2019). Does the use of summative peer assessment in collaborative group work inhibit good judgement? *Higher Education*, *77*(5), 853-870. https://doi.org/10.1007/s10734-018-0305-7

Clinkenbeard, P. (2012). Motivation and gifted students: Implications of theory and research. *Psychology in the Schools*, *49*(7), 622-630.

Cohen, E. G., & Lotan, R. A. (2014). *Designing groupwork: Strategies for the heterogeneous classroom* (3rd ed.). Teachers College Press.

Cohen, E. G., Lotan, R. A., Abram, P. L., Scarloss, B. A., & Schultz, S. E. (2002). Can groups learn? *Teachers College Record*, *104*(6), 1045-1068. https://doi.org/10.1111/1467-9620.00196

Cotic, M., Doz, D., Felda, D., Zakelj, A., & Zuljan, M. V. (2023). Teachers' appraisal of various aspects of the efficiency of homogeneous and heterogeneous learning groups. *International Electronic Journal of Mathematics Education*, *18*(1), 1-14.

De Hei, M., Admiraal, W., Sjoer, E., & Strijbos, J.-W. (2018). Group learning activities and perceived learning outcomes. *Studies in Higher Education*, *43*(12), 2354–2370. https://doi.org/10.1080/03075079.2017.1327518

Foley, R. C. (2012). *How differentiated groups affect fifth grade science students* [Unpublished master's thesis]. Montana State University.

Kanevsky, L. S., & Sheppard, S. (1999). Nurturing gifted students' metacognitive awareness: Effects of training in homogeneous and heterogeneous classes. *Roeper review*, *21*(4), 266-272.

Manfre, J. (2022, February 14). *The advantages of heterogeneous student groups in math*. Edutopia. https://www.edutopia.org/article/advantages-heterogeneous-student-groups-math

Johnson, R. T., & Johnson, D. W. (1994). An overview of cooperative learning. *Creativity and collaborative learning*. Brookes Press, Baltimore.

Lewis, B. (2019, July 7). *Heterogeneous definition in a classroom*. ThoughtCo. https://www.thoughtco.com/heterogeneous-groups-in-educational-settings-2081645

Lotan, R. A. (2003, March). Group-worthy tasks. *Educational Leadership*, *60*(6), 72-75.

Naperville central high school. (2021). Public School Review. https://www.publicschoolreview.com/naperville-central-high-school-profile#:~:text=Naperville%20Central%20High%20School%20serves%202%2C640%20students%20in%20grades%209%2D12

Naperville city, Illinois. (2021, July 1). United States Census Bureau. https://www.census.gov/quickfacts/napervillecityillinois

Schultz, S. E. (1999). To group or not to group: Effects of groupwork on students' declarative and procedural knowledge in science. [Unpublished dissertation]. Stanford University, Stanford, CA.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4

APPENDICES

APPENDIX A

POST UNIT SELF-REVIEW SURVEY

Please note that participation in this survey is completely voluntary. Your participation or nonparticipation as well as your answers on this survey will not impact on your grade or standing in this class in any way. Your specific answers will only be shared with Mr. Dille. Please answer as honestly as possible!

Last Name: _____________

First Name: _____________

During the last unit, how engaged were you during class time on a scale of 1-5, with 5 meaning you were almost always engaged in the physics content?

Almost never				Almost always
1	2	3	4	5

During the last unit, how often did you work with someone at your table to solve or discuss problem on a scale of 1-5, with 5 meaning you almost always worked with someone at your table?

Almost never				Almost always
1	2	3	4	5

For both of the responses above, briefly describe why you answered the way you did.

____

____

____

____

____

APPENDIX B

POST UNIT PEER-REVIEW SURVEY

Please note that participation in this survey is completely voluntary. Your participation or nonparticipation as well as your answers on this survey will not impact on your grade or standing in this class in any way. Your specific answers will only be shared with Mr. Dille. Please answer as honestly as possible!

Last Name: __________

First Name: __________

During **labs** in the last unit, how much of the group work did you personally contribute?

a. I contributed more than anyone else in the group.
b. I contributed about as much as the rest of my group.
c. I did not contribute as much as some of my peers.

During discussions or practice problems in the last unit where you discussed a topic or problem with your table group, which of the following best describes your role in the group?

a. I came up with more of the ideas or solutions than my peers.
b. I was actively engaged in the discussion, and we came up with ideas or solutions together.
c. Other members of my group came up with more of the ideas or solutions than me.

For both of the responses above, briefly describe why you answered the way you did.

____

____

____

____

____

How many members of your lab/table group were regularly **disengaged** during labs or group discussions (not including yourself)?

a. 0 (everyone was typically engaged)
b. 1
c. 2
d. 3

APPENDIX C

INSTITUTIONAL REVIEW BOARD APPROVAL

At 9:14 AM on Thursday, December 14, 2023, NoReply@TOPAZTI.com wrote:
Hello Dille, Nathanael,

Your protocol was reviewed by the IRB and has been approved.

PI: Dille, Nathanael
Approval Date: 12/14/2023
Title: Homogeneous and Heterogeneous Peer Groups in the Secondary Science Classroom

Protocol #: 2023-1004-EXEMPT
Review Type: Exemption
Expiration Date: 12/12/2028

Work described under this protocol may now commence. The PI is responsible for ensuring that the protocol accurately describes research practices being conducted.

> Review Category designation determined by the IRB can be found in the final section of your protocol.
> IRB-stamped active Consent Forms are attached within your protocol where applicable.
> Any changes must be submitted via Amendment prior to implementation.
> Per the Common Rule, research only requires Interim (annual) Review by the IRB if 1) it was reviewed via Full Committee or 2) is regulated by the FDA.
> All research is subject to post approval monitoring.
> All protocol types must be renewed 5 years after approval.
> Inform the IRB once your research is complete so that the protocol may be inactivated.

Please contact your IRB Program Manager with any questions or if you are in need of assistance. Thank you for your diligence in the care of human subjects research participants.

Institutional Review Board for the Protection of Human Subjects | Office of Research Compliance | Montana State University

APPENDIX D

PARTICIPANT ASSENT/CONSENT FORM

SUBJECT CONSENT FORM FOR PARTICIPATION IN HUMAN RESEARCH AT MONTANA STATE UNIVERSITY

Greetings parents, guardians, and students, as a part of my graduate work and professional learning, I am conducting a study with Montana State University on the effects of different grouping strategies on learning in the science classroom

You are being asked to participate in a research study pertaining to the effects of student lab and discussion groups on learning and engagement in a secondary science classroom. By placing students in several different student groups, recording subsequent scores on individual and group assignments, and surveying students about their experience in the groups, I hope to develop a better understanding of how group composition can improve the learning experience of individual students.

Your participation in this research is voluntary, and your participation or non-participation will not affect your grades or class standing in any way. If you agree to participate, you will be asked to complete a survey at the end of each unit. You may choose to not answer any question you do not want to answer and/or you may choose to stop participating at any time.

In addition to completing a survey at the end of each unit, students' scores on assessments, including tests, quizzes, and labs, will be used to measure the efficacy of different grouping arrangements. Note that all students will be working in student groups regardless of participation, as a regular element of a lab-based science course.

Risks: There are no foreseen risks beyond the normal risks associated with a lab-based science course.

Benefits: This study has no direct benefits to participants.

Source of funding: Not applicable. Cost to subject: None.

Confidentiality: Only the primary investigator/teacher, Nathanael Dille, will have access to identifiable student data, which will be stored on a district computer owned by Naperville CUSD 203 and operated by Nathanael Dille. While data from student scores and direct quotes from student surveys will be used in the final report, to be published to Montana State University's website, no student names or identifiable student information will be included in the report nor shared with any other person than the primary investigator/teacher.

If you have any questions regarding the research, you can contact Nathanael Dille at ndille@naperville203.org. If you have additional questions about the rights of human subjects, you may contact the Chair of the Institutional Review Board, Mark Quinn, at (406) 994-4707 [mquinn@montana.edu].

AUTHORIZATION: I have read the above and understand the inconveniences and risks of this study. I, ________________ (*name of parent or guardian*), related to the subject as _____________ (*relationship*), agree to the participation of ________________ (*name of subject*) in this research. I understand that the subject or I may later refuse participation in this research and that the subject, through her/his own action or mine, may withdraw from the research study at any time. I have received a copy of this consent form for my own records.

Parent or Guardian Signature: ______________________ Date: __________

Child's Assent Signature: ______________________

Investigator/Teacher: ______________________ Date: __________