2023 Annual Conference & Exposition
Baltimore Convention Center, MD | June 25 - 28, 2023
The Harbor of Engineering Education for 130 Years
Paper ID #39576

Unconventional Applications of Introductory-Level Aerospace Engineering Concepts: Evaluating Student Engagement and Performance in a Free-Response Exam Format

Benjamin Casillas, Texas A&M University

Ben Casillas is a senior aerospace engineering major at Texas A&M University. As an undergraduate researcher at the NUANCED Laboratory, their work focuses on novel presentations of introductory-level curriculum. Outside the lab, their interests include chemical rocket propulsion, spaceflight human systems integration, digital art, and music composition.

Dr. Kristi J. Shryock, Texas A&M University

Kristi J. Shryock, Ph.D., is the Frank and Jean Raymond Foundation Inc. Endowed Associate Professor in Multidisciplinary Engineering and Affiliated Faculty in Aerospace Engineering at Texas A&M University. She also serves as Director of the Novel Unconventional Aerospace Applications iN Core Educational Disciplines (NUA2NCED) Lab and of the Craig and Galen Brown Engineering Honors Program and National Academy of Engineering Grand Challenges Scholars Program. She has made extensive contributions to the methodology of forming the engineer of the future through her work in creating strategies to recruit, retain, and graduate engineering students. The network of transformational strategies she has developed addresses informing early, preparation for success, increasing diversity of the field, establishing strong identity as an engineer, and enhancing critical thinking and professional skills.

Unconventional Applications of Introductory-Level Aerospace Engineering Concepts: Evaluating Student Engagement and Performance in a Free-Response Exam Format

Abstract

Engineering is a broad field that covers a wide range of disciplines. Many institutions of higher education divide engineering into different departments, placing students into preset tracks with a particular focus. Still, many concepts are applicable across different fields, and skilled engineers develop an understanding of the interactions and similarities between disciplines. The NUA^{2}NCED Laboratory in Texas A&M's Department of Aerospace Engineering was established as part of an effort to encourage students to apply their knowledge beyond the traditional boundaries of aerospace engineering, and in so doing address a critical need for improved diversity within the department. This study applies introductory-level aerospace concepts in unconventional ways, while formatting them in traditional exam-format problems. The aim is to investigate whether this approach can be used to improve students' engagement in and understanding of the curriculum. AERO 201, which covers basic aerodynamic principles, was targeted for study. Students enrolled in this course were approached during a recitation period and introduced to the research. Students who chose to participate were divided into an experimental and a control group. Both groups completed a pre-survey assessing their current feelings about the course. Participants in the experimental group then completed a problem applying aerodynamic principles to a sailboat, while those in the experimental group completed a similar problem featuring a gliding aircraft. All participants then completed a second problem, which utilized similar concepts to the first problem but at a higher difficulty level. This problem, common to both groups, featured a climbing aircraft. Upon completion of both problems, participants completed a post-survey assessing their subjective feelings about the experiment. The results of the pre- and post-surveys are compared against each other and between the two groups to characterize any changes imparted by the experiment, as well as differences between the two groups. Additionally, participants' performance on both the first and second problem is assessed and compared between the two groups. Preliminary results suggest marginal effectiveness of this approach. While participants performed worse on the unconventional problem, they considered it just as relevant and helpful as a conventional problem. The results also highlight a need for adjusted measures of success in future investigations.

Introduction

The aerospace industry is constantly growing in its influence on the world. Developments over the past decade have focused on sustainable long-distance aviation technologies, urban air mobility, low-cost access to space, and the commercialization of human spaceflight, to name but a few. These and other projects continue to demand talented engineers to support their research and development. The NSF REDO-E grant supporting this study identifies several ways in which diversity in engineering consistently yields improvements across many facets of the discipline. Groups with higher diversity consistently demonstrate improved overall performance [1] as well as improved understanding of relevant subject matter [2]. Such groups are also more adept at making ethical decisions [3], which is of especially critical importance in the aerospace industry.

With these points in mind, the Department of Aerospace Engineering at Texas A&M University is but one of many in the United States that do not adequately reflect the diversity of its population as a whole. Women are heavily underrepresented as undergraduate students in this major, comprising just 8.3% of Bachelor's degrees awarded during the 2020-2021 academic year [4]. Ethnic minorities were similarly underrepresented during this academic year, with whites accounting for 67.5% of awarded Bachelor's degrees in aerospace engineering [4]. Enrollment figures reported by the university in fall of 2022 reflect slightly higher representation, with women comprising 14.1% of students in the aerospace engineering department, and whites comprising 56.03% of the enrolled population [5]. Texas is one of several states in which nationwide minority populations represent a majority of the local population, with non-hispanic whites comprising just 40.3% of the overall population [6]; Texas A&M's enrollment numbers do not reflect the diversity of its locale. As such, the NSF REDO-E grant identifies a significant deficiency in Texas A&M's ability to attract these individuals to aerospace engineering, jeopardizing its ability to address the needs of the industry.

This study seeks to address this issue by specifically targeting one aspect of the general concept of diversity: students with different backgrounds have different learning styles. These students may therefore vary in how attractive or effective they find the current state of aerospace curriculum to be. By implementing different approaches to engineering education, Texas A&M can enable more students from these underrepresented populations to become successful members of the industry.

In particular, this research focuses on AERO 201, the first introductory-level course in Texas A&M's aerospace program track. Students from this course are selected as participants in order to sample a population of individuals who are first encountering major-specific content in their education, and so may represent a more plastic demographic than students who are further along in the program. Additionally, should measures such as those demonstrated in this research be implemented in the future, this population of students represent those subject to the critical "first impression" which may determine their overall investment in the major. By presenting these students with unconventional applications of traditional aerospace engineering concepts, this study hopes to begin to characterize the effectiveness of this approach in improving student's engagement in and understanding of the material. The results of this and future related research may inform future improvements in engineering education seeking to improve diversity nationwide.

Background & Development

The nature and scope of this study was developed and revised over a period of three academic semesters. The concept began with exploratory development of several problems for students to complete, with similarly unconventional application of various engineering concepts. These problems were initially developed for use in a sketch recognition program. Difficulties with the implementation of this program, and the need for a more defined research purpose, led to a shift in the scope of the project. The sketch recognition program was replaced with a more traditional exam-format presentation, allowing increased focus on the unconventional content of the problems. The research presented here represents a smaller, more focused step towards understanding teaching methods in engineering.

The AERO 201 course at Texas A&M, which focuses on basic aerodynamics and some vector operations, was selected for use in this study. The basic premise was to develop problems with a similar format to those that students would encounter in AERO 201. These problems made use of similar engineering and physics concepts as those encountered in normal coursework, differing only in the object that these concepts were applied to. This is summarized in Table 1 below.

Table 1. Conventional vs. unconventional problem features

Problem feature	**Conventional problem**	**Unconventional problem**
Lift & drag equations	✓	✓
Interpreting physical diagrams	✓	✓
Vector component resolution	✓	✓
Algebra & trigonometry	✓	✓
Newton's second law	✓	✓
Application	**Aircraft**	**Sailboat**

In an informal trial, several AERO 201 students provided feedback on the design of the sailboat problem. This led to revisions to the problem statement and its diagram to improve clarity. Additionally, Dr. Kyle DeMars, an aerospace engineering professor at Texas A&M, was consulted for additional input. A rubric was developed by the author with the assistance of other lab members for Dr. DeMars to complete. The rubric is presented in Table 2; Dr. DeMars' selections are highlighted and bolded.

This and additional verbal feedback was used to refine the sailboat problem into its final state. The corresponding conventional version of the problem, as well as the second, more difficult problem, were then designed according to the input received regarding the sailboat problem. The final problems, as they were administered to students, are presented in Figures 1, 2, and 3.

Table 2. Rubric for problem feedback

Criteria	Ratings				
Clarity: Problem Statement Is the problem statement clear and easy to follow? Does it provide all information needed to complete the problem?	5 The problem statement is clear, easy to understand, and provides all necessary information to solve the problem.	**4** **The problem statement is mostly clear and provides helpful information. A few details may be unclear or confusing.**	3 The problem statement is somewhat clear, but some necessary information is missing or confusing.	2 The problem statement is difficult to understand, and most necessary information is missing or confusing.	1 The problem statement is nonsensical and provides no useful information.
Clarity: Picture Is the picture clear and easy to understand? Does it serve to further clarify the problem statement?	**5** **The picture is clear, easy to understand, and enhances the overall clarity of the problem.**	4 The picture is mostly clear and provides helpful insight. A few details may be unclear or confusing.	3 The picture is relevant to the problem, but somewhat confusing. It does not harm overall clarity.	2 The picture is mostly confusing, and decreases the overall clarity of the problem. It may contradict the problem statement.	1 The picture is nonsensical, provides no useful information, and contradicts the problem statement.
Difficulty Rank the problem difficulty based on the ratings given. This rating should be based on the perspective of the students.	5 The problem is unreasonably difficult and few students could attempt it. Most of the material is outside the scope of the class.	4 The problem is difficult and few students would complete the problem. Other students could attempt it but not complete it.	**3** **The problem is very well-suited to the skill level of its intended audience. It is neither trivial nor unreasonably difficult.**	2 The problem is not especially challenging, and a significant portion of students would find it trivial to complete.	1 The problem would be trivial to complete for most students. It would be possible to complete the problem without knowledge of the course concepts.
Concept Delivery Does the problem provide effective insight to the concepts taught in the course?	5 The problem is an intuitive representation of the concept and enhances students' understanding of the material.	**4** **The problem is a logical representation of the concepts provides some insight to the material presented.**	3 The problem is a logical representation of the concepts and provides little to no further insight to the material presented.	2 The problem is not a logical representation of the concept. It does not contribute to students' understanding of the material.	1 The problem is not a logical representation of the concept. It actively harms students' understanding of the material.

Problem 1: Glider

Although fixed-wing aircraft require airspeed to generate lift, they don't always need propulsion in order to fly. Gliding aircraft develop lift as they fall, allowing them to fly unpowered at some shallow angle from the horizontal.

a) Consider the glider below, which moves at speed v at an angle θ below the horizontal, shown as the +x direction. First, write an expression for the total force in the +x direction, F_x, as a function of θ, L, and D.

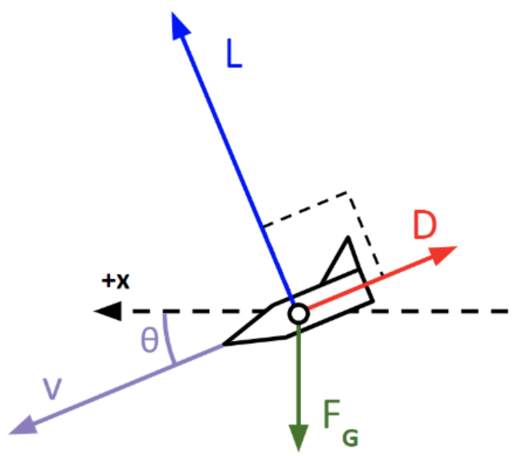

Figure 1: Forces on a glider

b) Next, use your expression to calculate the force F_x. The following parameters are provided:

v = 60 m/s θ = 30° ρ = 1.225 kg/m^3 C_L = 0.60 C_D = 0.02 S = 50 m^2

$L = ½ \rho v^2 S C_L$ $D = ½ \rho v^2 S C_D$

c) If the weight of the glider F_G = 5000 N, will the glider's total speed v increase, decrease, or stay the same?

Figure 1. Conventional first problem, administered to group A

Problem 1: Sailboat

While modern ships are predominantly concerned with hydrodynamics, historically aerodynamics was just as important in evaluating their performance. Sailing ships are able to travel upwind because their sails act as airfoils, producing lift at an angle that pulls them forward.

a) Consider the sailboat below, which encounters a wind speed v_w at an angle θ from its direction of travel, shown as the +x direction. First, write an expression for the total force in the +x direction, F_x, as a function of θ, L, and D. **Do not include the force K at this time.**

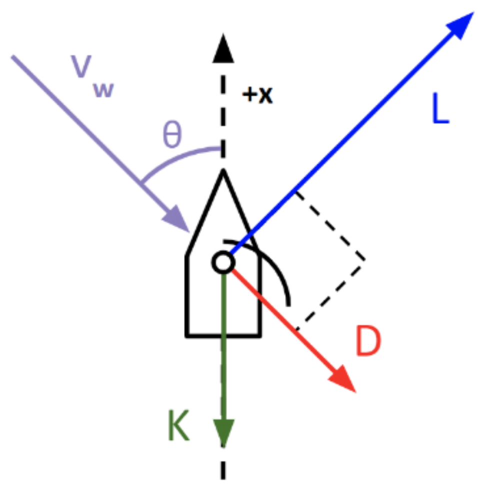

Figure 1: Forces on a sailboat

b) Next, use your expression to calculate the force F_x. The following parameters are provided:

$v_W = 6$ m/s θ = 30° $\rho = 1.225$ kg/m³ $C_L = 0.90$ $C_D = 0.01$ $S = 10$ m²

$L = \frac{1}{2}\rho v^2 S C_L$ $D = \frac{1}{2}\rho v^2 S C_D$

c) The water also exerts a separate drag force K on the boat. If this force K = 80 N, will the boat's speed increase, decrease, or stay the same?

Figure 2. Unconventional first problem, administered to group B

Problem 2: Climbing aircraft

a) Consider the aircraft below, which climbs at speed v at an angle θ above the horizontal, shown as the +x direction. The aircraft is powered by an engine producing a thrust force T. Using a similar approach to the first problem, write an expression for the net force in the +x direction as a function of θ, L, D, and T.

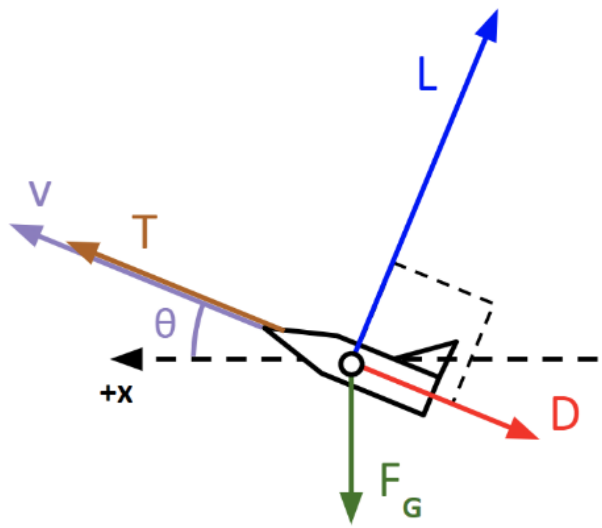

Figure 2: Forces on a climbing aircraft

b) Use the following parameters to calculate the net force in the +x direction.

$v = 180$ m/s $\theta = 15°$ $\rho = 1.225$ kg/m^3 $C_L = 0.50$ $C_D = 0.04$ $S = 35$ m^2

$T = 2500$ N $L = \frac{1}{2}\rho v^2 S C_L$ $D = \frac{1}{2}\rho v^2 S C_D$

c) If the weight of the aircraft $F_G = 3000$ N, will the aircraft's total speed v increase, decrease, or stay the same?

Figure 3. Second problem, common to both groups

Methods

Students were approached during a recitation period for AERO 201 and presented with a brief introduction to both the NUA^{2}NCED Laboratory and the experiment itself. The recitation period occurred from 9:10 AM to 10:00 AM in a lecture hall seating approximately 40 students. After the introduction, a total of 37 students opted to participate in the experiment.

The experiment as administered is divided into three segments: a pre-survey, two exam-format problems, and a post-survey. All materials were accessed online; participants were provided a series of internet links to direct them to each component of the study. Each participant was provided a single sheet of paper on which to write any scratch work necessary to complete the problems.

To protect students' identities, no confidential or personally identifiable information was collected, including previous grades or GPA information. However, each participant was expected to produce several distinct records: one response for each survey, an online submission covering both problems, and their scratch work. All of these documents needed to be associated with the participant that created them. To accomplish this, a double letter code system was created. The first letter of the code, either A or B, indicated whether the participant was in the control (A) or experimental (B) group. This distinction was not explained to the participants. The second letter of the code ranged from A to Z, accommodating up to 26 participants in each group. One of each double letter code (AA through AZ, BA through BZ) was printed in advance on each sheet of scratch paper. These scratch papers were randomly distributed to the participants, assigning each of them a unique code. Each online component of the experiment included a field for participants to enter their double letter code, allowing their records to be anonymously associated with each other. In all, 19 participants were assigned to group A, and 18 participants were assigned to group B.

Participants were first directed to complete a pre-survey ahead of their exposure to the two exam-format problems. Participants completed several optional demographics questions about their age, gender, and race/ethnicity, as well as a question identifying whether they had taken the course before, and the highest level of aerospace coursework they were enrolled in at the time of the experiment.

The pre-survey then asked participants to report their agreement with each of three statements on a Likert scale from 1 (Strongly Disagree) to 9 (Strongly Agree). These questions were used to provide insight into the mindset of students participating in the experiment. Each question was asked in both the pre-survey and the post-survey. This allowed for a characterization of students' engagement in and understanding of the material before and after completing the exam-format problems. These three statements are presented in Table 3 below.

Table 3. Pre-survey questions

Code	Question statement
PP1	I am confident that I want a career in Aerospace Engineering.
PP2	The material in AERO 201 is interesting to me.
PP3	The material in AERO 201 is difficult for me.

Participants were then directed to one of two online surveys containing the two exam-format problems, depending on the first letter of their double letter code. Only the first problem differed between the two groups; the control (A) group received the gliding aircraft problem, while the experimental (B) group received the sailboat problem. After completing the first problem, participants proceeded to the second problem, which was common between both groups and featured a climbing aircraft. Participants were instructed to record their scratch work on the paper provided, while their answers for both problems were submitted in the online survey form.

The second problem serves as a control for participants' skill level. Additionally, the act of completing the first problem may influence students' approach to the second problem, and this influence may differ between participants who receive the conventional and unconventional versions of the first problem. Although this does not fully eliminate uncertainty regarding the effectiveness of the unconventional first problem, it provides additional data that can be used to glean more information about this process. A third problem, to be completed prior to the others and thus serve purely as an accurate control for participants' skill level, was considered but rejected so that participants could complete the experiment in a reasonable amount of time. Instead, the pre-survey asked participants to self-report the difficulty of the course content.

Upon completion of both problems, all participants were directed to complete a post-survey. This survey contained no demographics questions. Participants were once again asked to report their agreement with questions PP1, PP2, and PP3. This was intended to provide insight into any effects the experiment may have had on participants, and how those effects may have differed between the two groups. To further elaborate on students' impressions of the problems, five retrospective questions were provided, which the participants responded to using an identical Likert scale as before. One optional question (R4) asked participants to provide a short-answer response. All of these questions were intended to help understand the effectiveness of the unconventional first problem compared to the conventional first problem. These questions are presented in Table 4 below.

Table 4. Retrospective questions

Code	Question statement
R1	The first problem made use of the skills I've learned in AERO 201.
R2	The second problem made use of the skills I've learned in AERO 201.
R3	The first problem influenced my approach to the second problem.
R4*	If you agreed with the previous statement, how did the first problem influence your approach to the second problem? (*Optional short-answer)
R5	The experience of completing the first problem was helpful in completing the second problem.
R6	I would be comfortable seeing a problem like the first one on a graded assignment in AERO 201.

After submitting both surveys, their responses to the exam-format problems, and their scratch paper, students' participation in the study was complete.

Results

19 usable responses were received in group A, and 16 usable responses were received in group B. The overall ethnic and gender demographics of the study population are described in Figures 4 and 5, respectively. Note that the race/ethnicity questions allowed multiple responses.

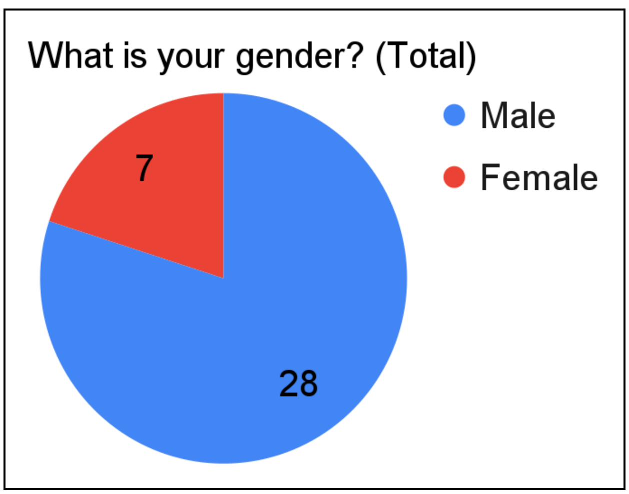

Figure 4. Participant gender responses

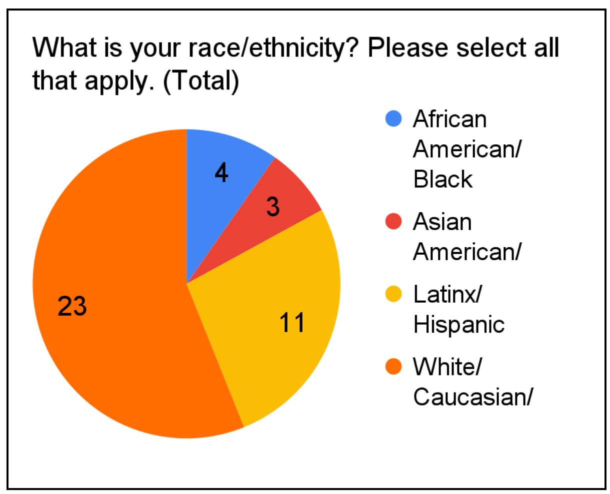

Figure 5. Participant race/ethnicity responses

The participant sample size slightly overrepresents women in the major with respect to recent enrollment figures; 20.0% (7/35) of participants reported their gender as female. In reporting their race/ethnicity, 51.4% (18/35) of participants selected only White/Caucasian/European American, with 65.7% selecting White/Caucasian/European American in total.

Participants were randomly assigned to either the control group (A) or the experimental group (B), as denoted by the first letter of the double letter code marked on the scratch paper they received. Gender demographics for the two groups are described in Figures 6 and 7.

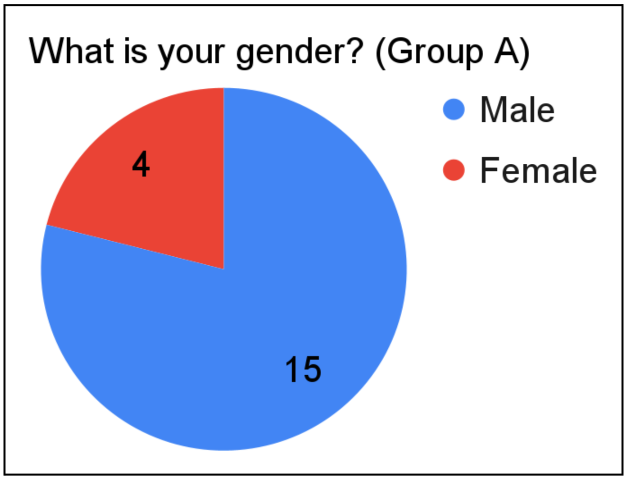

Figure 6. Group A gender responses

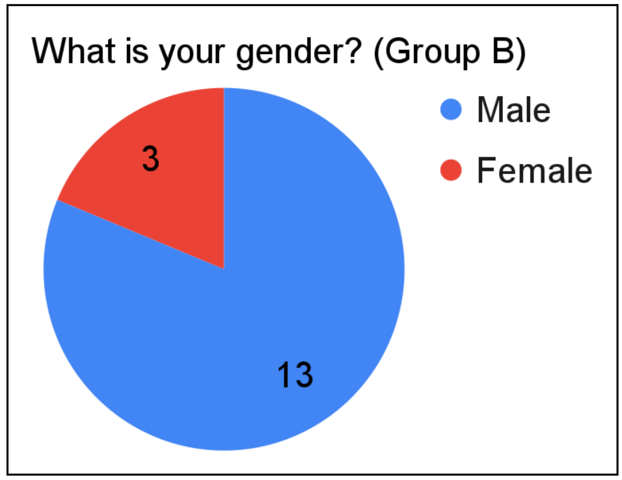

Figure 7. Group B gender responses

Gender representation varies little between the two groups, with both being fairly representative of the overall sample population. The racial/ethnic demographics of the two groups are described in Figures 8 and 9.

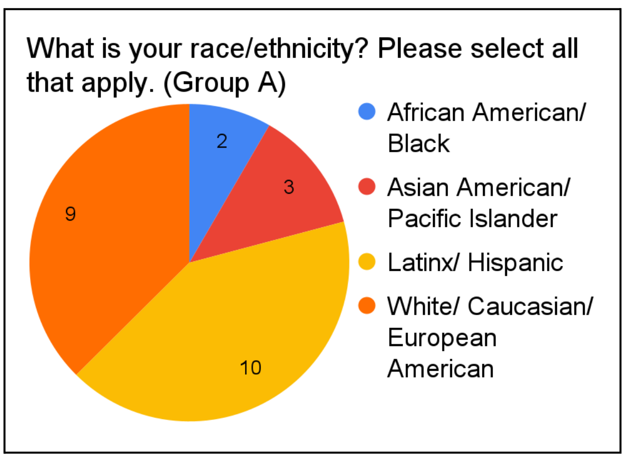

Figure 8. Group A race/ethnicity responses

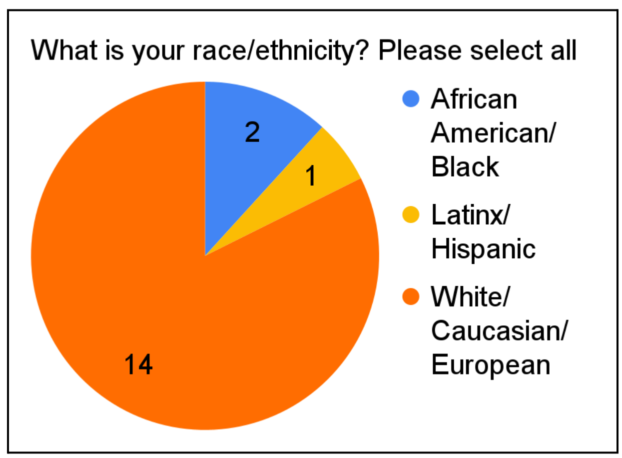

Figure 9. Group B race/ethnicity responses

Notably, the two randomly assigned participant groups are radically different in their ethnic makeup. Group A reflects a "majority minority" environment in which participants who selected only White/Caucasian/European American represent only 26.3% (5/19) of the group. Conversely, group B is overwhelmingly composed of white individuals, with 81.3% (13/16) selecting only White/Caucasian/European American. Neither group is closely representative of the overall study population.

Three of the Likert scale questions, PP1, PP2, and PP3, were common to both the pre- and post-survey, allowing a comparison between participants' responses before and after their exposure to the problem sets. Table 3 provides descriptions of these questions. Table 5 contains a summary of relevant data for all three questions, including both the pre-survey and post-survey responses, for Groups A and B.

Table 5. Pre- and post-survey comparison responses

Question	Value	Pre Survey	Post Survey	Delta	Delta (%)
PP1A	Mean	7.632	7.579	-0.053	-0.690
	SD	1.707	1.742	0.036	2.087
PP2A	Mean	8.316	8.263	-0.053	-0.633
	SD	1.250	1.195	-0.055	-4.404
PP3A	Mean	4.947	5.211	0.263	5.319
	SD	1.545	1.512	-0.033	-2.105
PP1B	Mean	8.563	8.563	0.000	0.000
	SD	0.727	0.727	0.000	0.000
PP2B	Mean	8.438	8.313	-0.125	-1.481
	SD	0.964	0.946	-0.017	-1.810
PP3B	Mean	5.125	5.250	0.125	2.439
	SD	1.857	1.807	-0.050	-2.693

The mean responses, compared between the pre- and post-surveys, are visually represented in Figure 10 below.

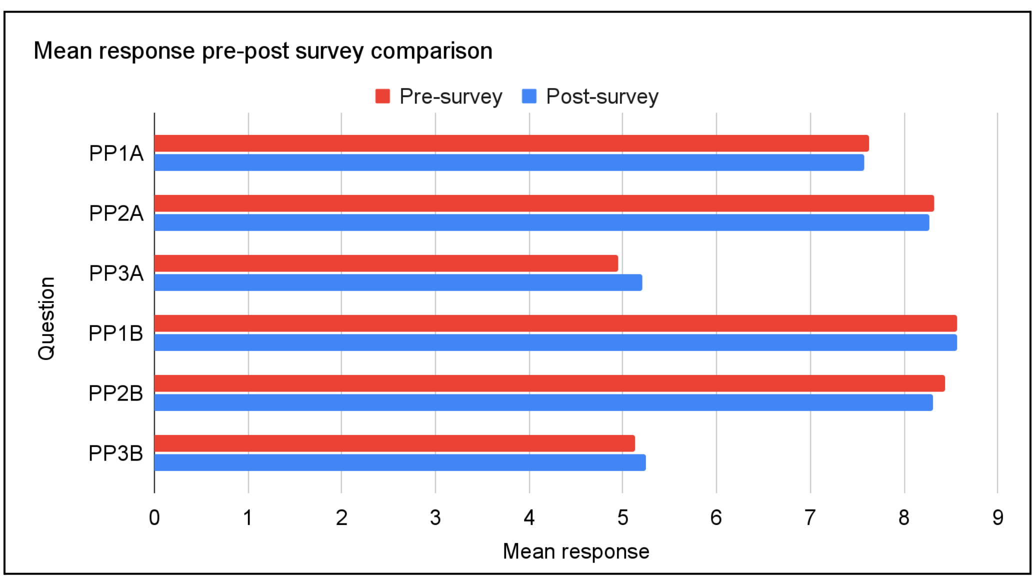

Figure 10. Mean response pre-post survey comparison

To investigate the significance of this data, a series of two-tailed paired t-tests were performed. For each question, and in each group, participants' responses in the pre-survey were compared with those in the post-survey to determine whether a significant change occurred. The results of this analysis are presented in Table 6. Note that the "df " column represents the number of degrees of freedom for the test.

Table 6. Two-tailed paired t-test results

Test	df	t	p
PP1A	18	0.567	0.578
PP2A	18	0.369	0.716
PP3A	18	-2.041	0.056
PP1B	15	-	-
PP2B	15	1.464	0.164
PP3B	15	-0.808	0.432

Participant responses for question PP1 in group B (denoted PP1B) were identical in both the pre- and post-surveys, so the t-test could not be performed. Using a significance criterion of $p < 0.05$, none of these differences are statistically significant. The experiment does not appear to have strongly impacted participants' opinions regarding each of the three statements which were common to both the pre- and post-survey. Perhaps the closest is PP3A, the question regarding students' perceived difficulty of AERO 201. This agrees with the unusually large change in the mean for this question, as depicted in Table 5.

These responses can also be compared between groups A and B. Figure 11 represents only the mean responses from the pre-survey, as the pre/post change was negligible.

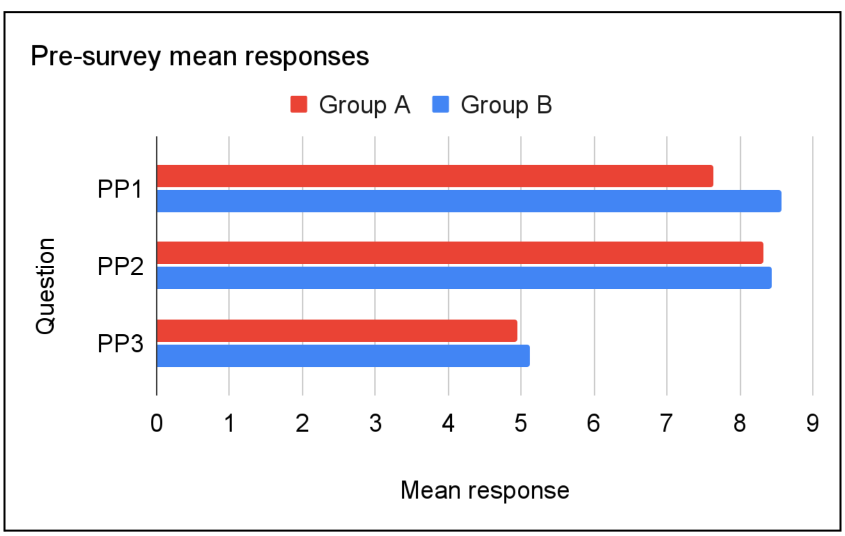

Figure 11. Pre-survey mean responses

A series of two-tailed independent t-tests was performed to characterize the differences between groups A and B. For each question, the pre-survey responses for group A were compared to those of group B. The results of this analysis are presented in Table 7.

Table 7. Two-tailed independent t-test results

Test	df	t	p
PP1AB	33	-2.029	0.051
PP2AB	33	-0.318	0.753
PP3AB	33	-0.309	0.759

Using a significance criterion of $p < 0.05$, none of these differences are statistically significant. The closest case is PP1, regarding students' confidence in their major, where the mean in group B is apparently, if not significantly, higher than that of group A. Thus, both groups had similar opinions on the interest and difficulty of the course, differing most noticeably in their overall confidence in their choice of major.

Five additional Likert scale statements were provided in the post-survey for students to submit opinions regarding their retroactive analysis of the experiment. Table 4 provides descriptions of each of these questions. The mean and standard deviations of responses for these questions, in groups A and B, are summarized in Table 8.

Table 8. Retrospective responses

Question	Value	Group A	Group B
R1	Mean	6.789	5.813
	SD	1.873	2.007
R2	Mean	7.158	6.313
	SD	1.500	1.922
R3	Mean	7.789	6.500
	SD	1.475	2.898
R5	Mean	7.316	6.625
	SD	1.600	2.187
R6	Mean	7.000	7.375
	SD	2.160	1.821

These results are depicted visually in Figure 12 below.

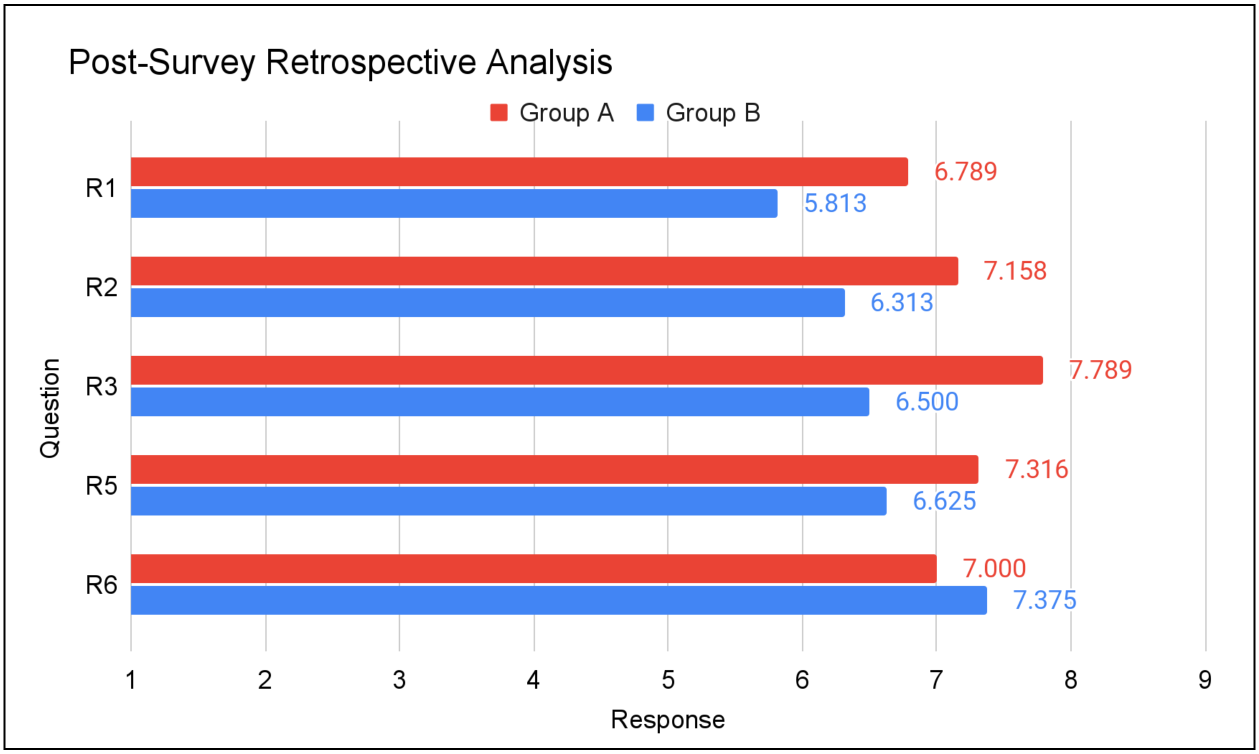

Figure 12. Post-survey retrospective analysis

To investigate the significance of this data, a series of two-tailed independent t-tests were performed. For each question, responses in group A were compared with those in group B to determine whether a meaningful difference existed. The results of this analysis are presented in Table 9.

Table 9. Two-tailed independent t-test results

Test	**df**	**t**	**p**
R1AB	33	1.488	0.146
R2AB	33	1.461	0.153
R3AB	33	1.7	0.099
R5AB	33	1.077	0.289
R6AB	33	-0.549	0.587

None of these differences are statistically significant with a criterion of $p < 0.05$. Although the differences in the mean appear substantial, the variance is high enough that this is attributable to random error. The R3AB test, if taken as a one-tailed result, yields $p = 0.049$, possibly indicating that the R3 response in group A is significantly higher than that in group B.

The author noted a significant discrepancy throughout the experiment in the time taken for participants to complete their problem sets if they were in group A as opposed to group B.

Although this analysis was unplanned, each response was timestamped, enabling this effect to be quantified. The results are presented in Figure 13; note that the totals here include participants whose responses were unusable for other reasons.

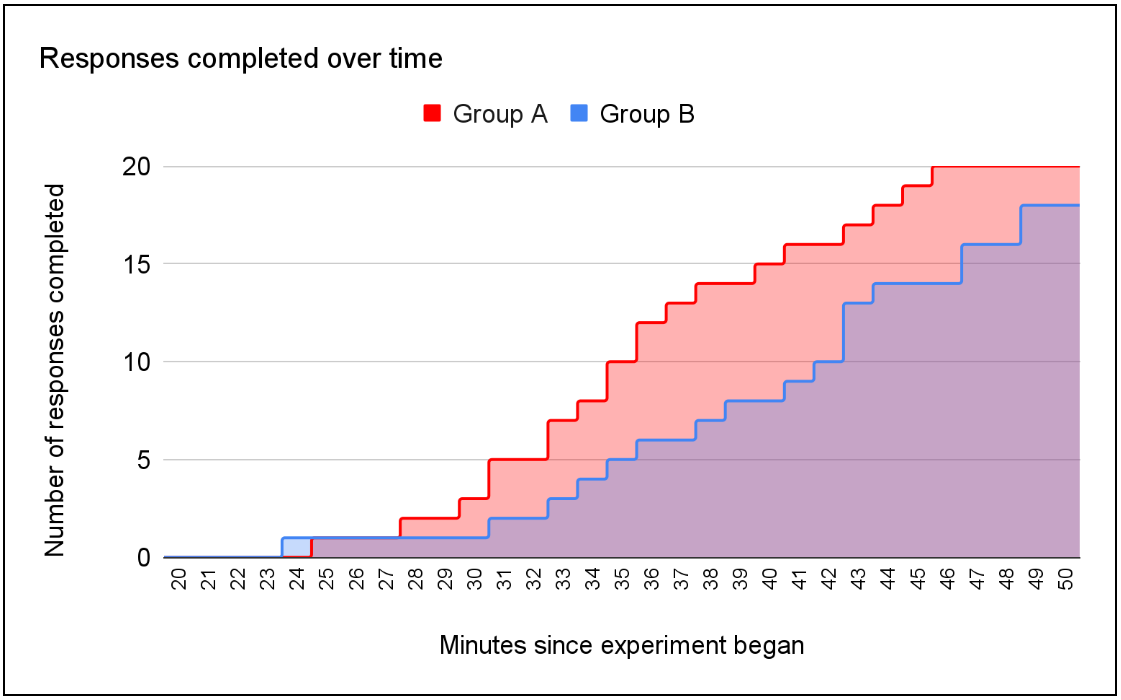

Figure 13. Responses completed over time

Participants in group A, who received a conventional first problem, consistently completed their responses earlier than those in group B, who received an unconventional first problem. Additionally, participants in group A completed their responses at a higher rate than those in group B. Using the last minute in which zero responses are complete in each group as a starting point, and ending at the minute of the final response's completion, group A shows completion at a rate of 0.909 responses per minute; group B's rate is 0.692 responses per minute.

These results appear to indicate that the unconventional problem required a longer time for students to complete than the conventional problem. This may be due to the unfamiliar nature of the problem when compared to problems traditionally presented in the course. Additionally, the response times in group B are shifted slightly to the right. The average completion time in group B was 39.89 minutes, compared to 36.05 minutes in group A; participants in group B completed their responses 3.84 minutes later on average than those in group A. Where the lower completion rate over time may indicate a greater difficulty for the unconventional problem, the consistent lag in completion times may represent a larger initial barrier in understanding, requiring participants to spend longer on initial comprehension before completing the problem.

All data analysis discussed to this point has concerned the Likert scale responses from the pre- and post-surveys. Participants also submitted their solutions to the exam-format problems.

Correct solutions to the problems were prepared, and participants' solutions were compared against these. Numerical answers were permitted a 5% margin of error if participants had used the correct equations to obtain their answers. Participants showed significant variation in the exact formatting of their equations, and there were multiple "correct" ways to frame the solution. Each part (a, b, and c) of each problem (1 and 2) was worth 1 point, for a maximum total of 6 points. The means and standard deviations of students' scores are summarized in Table 10 below. Note that S1 refers to students' score on problem one, while S2 refers to students' score on problem 2.

Table 10. Student score summary

Problem	Value	Group A	Group B
S1	Mean	1.895	1.063
	SD	1.197	1.124
S2	Mean	1.684	1.688
	SD	1.108	0.873
Total	Mean	3.579	2.750
	SD	2.143	1.612

The mean scores are visually represented in Figure 14.

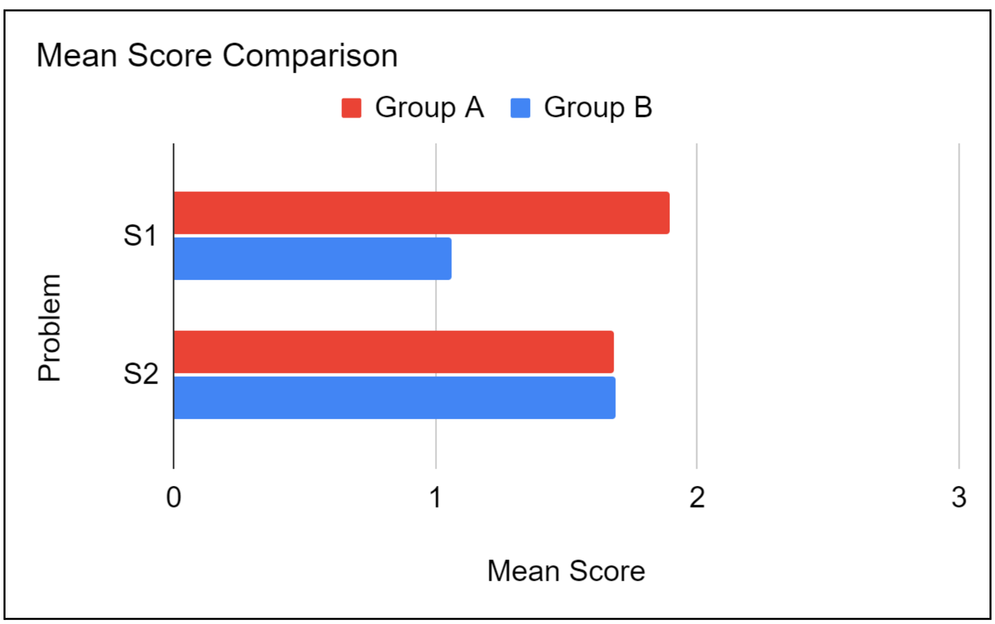

Figure 14. Participant mean scores

A striking difference is evident between the S1 means in groups A and B. Another set of two-tailed independent t-tests was performed to test for statistical significance. These results are presented in Table 11 below. Note that S3 refers to the difference between S2 and S1 scores.

Table 11. Two-tailed independent t-test results

Test	df	t	p
S1AB	33	2.107	0.043
S2AB	33	-0.010	0.992
S3AB	33	-2.394	0.023

With a significance criterion of $p < 0.05$, both S1AB and S3AB results were significant. This indicates that participants in group B scored significantly lower on problem 1 than those in group A. Additionally, the difference in participants' scores between problems 1 and 2 was significantly greater in group B than in group A.

Finally, for each group, a correlation study was performed between participants' scores on the two exam-format problems (S1 and S2), and their responses to each of the retrospective questions (R1, R2, R3, R5, and R6). This was used to further investigate more nuanced relationships and how they differed between the two groups.

These relationships, evaluated in group A, are presented as a correlation matrix in Table 12.

Table 12. Correlations in group A

	S1	S2	R1	R2	R3	R5	R6
S1	1.000						
S2	0.727	1.000					
R1	0.485	0.662	1.000				
R2	0.350	0.600	0.665	1.000			
R3	0.207	-0.009	-0.178	0.041	1.000		
R5	0.395	0.216	0.153	0.395	0.642	1.000	
R6	0.236	0.487	0.577	0.531	0.244	0.578	1.000

The strongest correlation in this group is between participants' scores on problems 1 and 2 (S1×S2), demonstrating the relatively consistent values in these two areas. Other relatively strong correlations include R1×R2 and R3×R5. These indicate that participants in group A felt both problems were similarly relevant to 201. If they felt that the first problem influenced their approach to the second problem, they considered this to be helpful. This is consistent with their condition as a control group, in which both problems were intended to represent conventional AERO 201 content. The weaker correlations are also of note. The low values for S2×R3 and S2×R5 indicate that a participant's score on the second problem does not strongly relate to whether they considered the first problem to be helpful or not. Also, the low value for S1×R6 shows little relationship between a student's performance on the first problem and the degree to which they'd be comfortable completing it in class.

The same relationships, evaluated in group B, are presented in a similar format in Table 13.

Table 13. Correlations in group B

	S1	S2	R1	R2	R3	R5	R6
S1	1.000						
S2	0.293	1.000					
R1	0.331	-0.150	1.000				
R2	0.176	-0.295	0.845	1.000			
R3	-0.194	-0.145	0.418	0.425	1.000		
R5	0.092	0.004	0.514	0.474	0.873	1.000	
R6	0.346	-0.005	0.403	0.440	0.366	0.540	1.000

The strongest correlation in this group is R3×R5. Participants in group B who felt that the first problem influenced their approach to the second problem consistently felt that this was a positive change. If it did not influence their approach, they considered it unhelpful. Interestingly, R1×R2 is another very strong correlation. This indicates that participants in group B also felt both problems were equally relevant to AERO 201, despite the first being presented unconventionally.

As with group A, participants in group B showed low values for S2×R3 and S2×R5. This indicates that there was not a strong relationship between participants' scores on the second problem and their opinion on whether the problem was helpful or not. Additionally, a fairly low value for S1×R6 again suggests that a student's performance on the first problem was unrelated to whether they'd be comfortable completing it in class.

Conclusions & Lessons Learned

The mission of the grant sponsoring this research focuses on addressing shortcomings in diversity in aerospace engineering. One solution proposed by the grant was to study the effectiveness of alternate or untried teaching methods in order to better reach populations such as women and minorities which have historically been underrepresented in the major. This research studied one possible implementation of this approach, utilizing an unconventional framing of familiar aerospace concepts in a relatively standard exam-format presentation. The major goals of the study were to characterize the effect of this method on students' engagement in and understanding of the course material, with the hope that this data may inform future improvements to aerospace engineering curriculum.

Regardless of their group, students' interest in the course, and its perceived difficulty, were unchanged by the experiment. The unconventional framing of aerospace material took place in a conventionally-formatted exam problem, which comprised only half of the overall experiment for students in the experimental group. The effects of this experimental design against the status quo could be expected to be small as a result.

The students' overall opinions about the course were also relatively homogeneous. Regardless of group, participants felt confident in their choice of major, and felt the content in AERO 201 was interesting. Participants also felt that AERO 201 was of moderate difficulty; the average response was slightly above the neutral "neither agree nor disagree" option. When participants' well-matched scores on the common problem 2 are also considered, the overall skill level of participants in both groups seems to be fairly even.

The most statistically significant result was the difference in scores between groups A and B. Students scored significantly worse on the unconventional problem 1, involving a sailboat, than they did on the conventional problem 1, involving an aircraft. Together with the information collected about their skill level, the result seems to be that the sailboat problem was more difficult for participants to complete than a more conventional problem. Since scores on the second problem were consistent between groups, the sailboat problem does not appear to have benefited participants who completed it. The higher completion times in group B may point to increased engagement with the material, as well as to higher difficulty or complexity of the sailboat problem.

In retrospect, this is a consequence of the overall experiment design which was not anticipated during development. When the success of an exam problem is gauged by its ability to directly prepare students for a follow-up problem, unfamiliar content framing seems to be at a natural disadvantage. Rather than an "apples-to-apples" comparison, this sort of research seems to demand an "apples-to-oranges" approach, where measures of success will need to be adjusted to accommodate natural differences in the material presentation.

However, the retrospective questions add an interesting layer to the situation. Despite students scoring objectively lower on the sailboat problem, participant responses to the retrospective questions were fairly consistent. Participants who received the unconventional problem did not feel that it was significantly less relevant to AERO 201, or that it was significantly more or less helpful in completing the second problem. Students in both groups considered any influence the first problem had on their approach to the second problem to be beneficial.

Overall, these results suggest that this unconventional sailboat problem does not meaningfully impact students' feelings about the course, including their engagement in or understanding of the material. The sailboat problem appears to be a worse fit for students' abilities than the conventional problems they are used to, as evidenced by the lower values for S1B. However, in spite of their lower performance, students in group B did not feel that the unconventional problem was detrimental to their learning. In fact, they felt just as strongly as their peers in group A that the unconventional problem was relevant to their education, and that it helped them to complete the second problem.

These results highlight a need to carefully consider success criteria in this approach to engineering education. Novel curriculum, by nature, is not intended to exactly replicate the status quo, and this will be reflected in attempts to measure its effectiveness. Nevertheless, students appeared receptive to unconventional material regardless of whether they performed well on it initially, which lends merit to the overall premise. Longer-duration studies may be needed to more effectively gauge the impact of these learning methods on students' educations.

While analysis of demographic effects is beyond the scope of this study, the ethnic makeup of the two participant groups demonstrates a hazard of random assignment in a small sample size. With high asymmetry between the two participant groups, additional uncertainty is woven into the results of the experiment. Future research into the effectiveness of novel curriculum in reaching these minority populations would do well to control for this effect.

Finally, this study collected a wealth of data, only some of which could be analyzed within the time constraints for this study. In addition to conducting new research, potential exists for forward work to simply perform additional analysis of the data collected during this experiment.

Acknowledgements

The author would like to formally acknowledge the REDO-E grant, project number 1704599, received from the National Science Foundation, without which this study would not have been possible, and whose mission motivated the author's desire to contribute to the development of engineering curriculum in pursuit of a representative and diverse future for the industry. The author would also like to thank Dr. Kyle DeMars for his insightful feedback regarding the problem sets; Hannah Stroud, a member of the NUA^{2}NCED Laboratory, for her invaluable guidance throughout the development of this study, and Dr. Kristi Shryock, for her assistance in coordinating the experiment itself, and for her indispensable mentorship and encouragement every step of the way.

References

[1] D. A. Harrison, K. H. Price, J. H. Gavin, and A. T. Florey, “Time, teams, and task performance: Changing effects of surface- and deep-level diversity on group functioning,” *Academy of Management Journal*, vol. 45, no. 5, pp. 1029–1045, 2002, https://doi.org/10.2307/3069328

[2] D. van Knippenberg, C. K. W. De Dreu, and A. C. Homan, “Work Group Diversity and Group Performance: An Integrative Model and Research Agenda,” *Journal of Applied Psychology*, vol. 89, no. 6, pp. 1008–1022, 2004, https://doi.org/10.1037/0021-9010.89.6.1008

[3] S. W. DeGrassi, W. B. Morgan, S. S. Walker, Y. Wang, and I. Sabat, “Ethical Decision-Making: Group Diversity Holds the Key,” *Journal of Leadership, Accountability and Ethics*, vol. 9, no. 6, pp. 51–65, 2012.

[4] “The Aerospace Engineering Major at Texas A&M University.” collegefactual.com. https://www.collegefactual.com/colleges/texas-a-and-m-university-college-station/academic-life/academic-majors/engineering/aerospace-and-aeronautical-engineering/ (accessed Feb. 20, 2023).

[5] “Accountability - Student Demographics.” accountability.tamu.edu. https://accountability.tamu.edu/all-metrics/mixed-metrics/student-demographics (accessed Feb. 18, 2023).

[6] “U.S. Census Bureau QuickFacts: Texas.” census.gov. https://www.census.gov/quickfacts/TX (accessed Feb. 25, 2023).