The Future of Engineering Education
2024 Annual Conference & Exposition
Oregon Convention Center
Portland, OR . June 23 - 26, 2024
ASEE
Paper ID #41965

Equation Sheets: Are We Helping or Hurting our Students?

Lt. Col. William Graves, United States Military Academy

William Graves, PhD, PE is an Assistant Professor in the Department of Civil and Mechanical Engineering at the United States Military Academy, West Point, New York.

Dr. Gary A Jordan, United States Military Academy

2018 PhD Civil Engineering (Transportation Systems Engineering)- SUNY Buffalo 2017 MS Civil Engineering - SUNY Buffalo 2012 MA Economics - University of New Hampshire 2011 MBA - University of Maine 1987 BS Engineering Physics - United States Military

Equation Sheets:
Are We Helping or Hurting our Students?

A common dilemma when administering significant graded events is whether to allow students to use a self-generated equation sheet, especially for closed-book examinations. While much research has evaluated the efficacy of these equation sheets on student performance, there remains a gap in understanding the effects of the quality of the equation sheets on exam performance. This study analyzes the relationship between and quantitative evaluation of student generated equations sheets and exam performance on a particular graded event in a course in a civil engineering program of study.

Students at the United States Military Academy take CE450, Construction Management, typically during their junior or senior year (of note, less than 1% of the students in the course are enrolled as sophomores). The distribution of academic majors among the students in the course is a mix between civil engineering, other STEM, and a variety of non-engineering majors. This mixture of academic major representation in the course is a result of an institutional graduation requirement that all non-engineering major students take a 3-course engineering sequence of classes. CE450, Construction Management, is one of the courses in the infrastructure engineering sequence of classes for the other STEM and non-engineering major students, and it is a required course for the civil engineering majors.

In the course final examination, students are allowed to bring a double-sided 8.5” x 11”, self-generated and hand-written note sheet as an aid. The requirement for the notes to be hand-written is intended to prevent cutting and pasting images or text from an outside source and to compel the students to solidify their understanding of the associated lesson objectives through transcription of key material. Equation sheets are collected at the end of the examination and evaluated to determine the associated density of the material recorded on the page. This word density is then used as a quantifiable metric to relate to exam performance.

Data was collected over the course of two sequential semesters in an academic year with a study population of 210 students. This study will be of interest to faculty and administrators that have an interest in or debate concerning the use of student-generated equation sheets in their respective courses.

Introduction

At the United States Military Academy, students are required to take a three-course engineering sequence (CES) to support their academic interests [1]. These three courses are mandatory for those students that do not major in one of the engineering disciplines, while the courses are otherwise incorporated into the academic programs for those students majoring in engineering. The options for the CES include infrastructure, cyber, robotics, environmental, nuclear, and systems engineering. The purpose behind the inclusion of these courses is to support several of the institutional level Academic Program Goals (APGs), which are further described via “What

Graduates Can Do" (WGCD) statements [2]. Table 1 lists the seven APGs (along with their descriptions) for the institution and the associated WGCD statements supported by the CES. For brevity purposes, WGCD statements supported by other aspects of the academic program are not included.

Table 1. Academic Program Goals and Associated "What Graduates Can Do" Statements for Three-course Engineering Sequence [2]

1. Communication: Graduates communicate effectively with all audiences
2. Critical Thinking and Creativity: Graduates think critically and creatively
2.1 Identify the essential aspects of a situation and ask relevant questions.
2.4 Reason both quantitatively and qualitatively.
2.5 Think innovatively and accept risk to pursue solutions in the face of ambiguity.
3. Lifelong Learning: Graduates demonstrate the capability and desire to pursue progressive and continued intellectual development.
4. Ethical Reasoning: Graduates recognize ethical issues and apply ethical perspectives and concepts in decision making.
4.2 Recognize ethical components of problems and situations.
5. Science/Technology/Engineering/Mathematics (STEM): Graduates apply science, technology, engineering, and mathematics concepts and processes to solve complex problems.
5.1 Apply mathematics, science, and computing to model devices, systems, processes, or behaviors.
5.4 Apply an engineering design process to create effective and adaptable solutions.
6. Humanities and Social Sciences: Graduates apply concepts from the humanities and social sciences to understand and analyze the human condition.
7. Disciplinary Depth: Graduates integrate and apply knowledge and methodological approaches gained through in-depth study of an academic discipline.
7.4 Synthesize knowledge and concepts from across their chosen disciplines.
7.5 Contribute disciplinary knowledge and skills as a part of a collaborative effort engaging challenges that span multiple disciplines.

The CES serves as a vital contributor to the institution reaching its APGs, and the focal point for the data in this research effort comes from one of the courses in the infrastructure CES. This sequence is intended to teach the students about the design, analysis, and construction of the built environment [1]. The courses are comprised of *MC300: Fundamentals of Engineering Mechanics and Design*, focusing on introductory statics and mechanics of materials; *CE350: Infrastructure Engineering*, teaching fundamentals related to major infrastructure sectors including water, power, transportation, waste management, etc.; and *CE450: Construction Management*, providing instruction on how to plan and execute the construction of the previously studied infrastructure components and systems. Typically, students take these three courses in the order listed. This approach ensures that the progression of information follows the original design and intent of the sequence, but exceptions are made to accommodate situations such as study abroad opportunities and scheduling irregularities.

Within the infrastructure CES, various courses allow the use of student-generated equation sheets for support on examinations. To understand the efficacy of self-generated equation sheets, this work evaluates the relationship in one of the classes between the performance on a major graded event and the average word density of the associated student-generated equation sheet. To that end, the authors have developed two basic research questions:

1. To what degree is there a relationship between the amount of content that a student places on a self-generated equation sheet and the performance on the associated graded event?
2. Do the relationships between the equation sheet and graded event reveal anything about our approach to teaching and evaluation?

Literature Review and Contributions of this Work

Research related to the efficacy of equation sheets (aka cheat sheets, crib sheets, calculation sheets, notecards, etc.) is mixed. To simplify terminology, this paper will refer to all forms of student produced notes approved for personal use during exams as equation sheets. Settlage and Wollscheid [3], Larwin, Gorman, and Larwin [4], Dickson and Bauer [5], and Song and Thuente [6] found equation sheets are associated with improved performance, while findings from Gharib, Phillips, and Mathew [7], and Sanborn, Purchase, and Barry [8] suggest equation sheets have mixed to no discernible impact upon performance. Some researchers, such as Vessey and Woodbury [9] argue that equation sheets may hurt performance by diverting student time, attention, and effort.

With respect to equation sheet attributes, researchers generally focus attention upon characteristics such as size, organization, and content. Danielian and Buswell [10] developed a four-tier rating system ranging from sparse to extremely dense. While they found increased word density was associated with an increase in course performance, it did not have the same effect upon exam performance – as word density increased, exam performance decreased. In addition to equation sheet characteristics discussed previously, researchers such as Shaw and Almeida [11], Hamouda and Shaffer [12], and Cannonier and Smith [13] have shown how specific disciplines vary in their use of equation sheets.

However, the authors' study focuses upon two specific attributes – one which is a characteristic of a given student's equation sheet, and the other which identifies the relationship between a student's course of study and the course type (or discipline) in which the equation sheet is used. Here again, our study provides a unique perspective. As previously noted, our course is part of a three-course engineering sequence in which students from all majors and disciplines may and do enroll. Because our student population within the course is heterogeneous, our study provides insight into how equation sheets may impact performance among and across academic disciplines.

Methods

As the students are allowed to develop their own individual equation sheets, there is variability in the amount of information that each individual student includes. The information on each equation sheet is quantified via the metric of its average word density. Average word density is calculated by sampling various locations on the equation sheet and taking the average number of words per given sampling area (further discussed in the next section).

Normalized grades are plotted against normalized average word density on the equation sheets. In each case, the normalized values are calculated by dividing the deviation from the mean value divided by the standard deviation of the data set, as seen by equation 1.

$$\hat{x} = \frac{x - \mu}{\sigma} \qquad \textit{(EQN 1)}$$

where

$$\hat{x} \equiv normalized\ value$$

$$x \equiv sample\ value$$

$$\mu \equiv population\ mean$$

$$\sigma \equiv population\ standard\ deviation$$

The arrangement of data in this manner results in four broad classifications based on the quadrant of the graph: a high grade supported by a dense equation sheet in Quadrant 1, a low grade even though it is supported by a dense equation sheet in Quadrant 2, a low grade with a

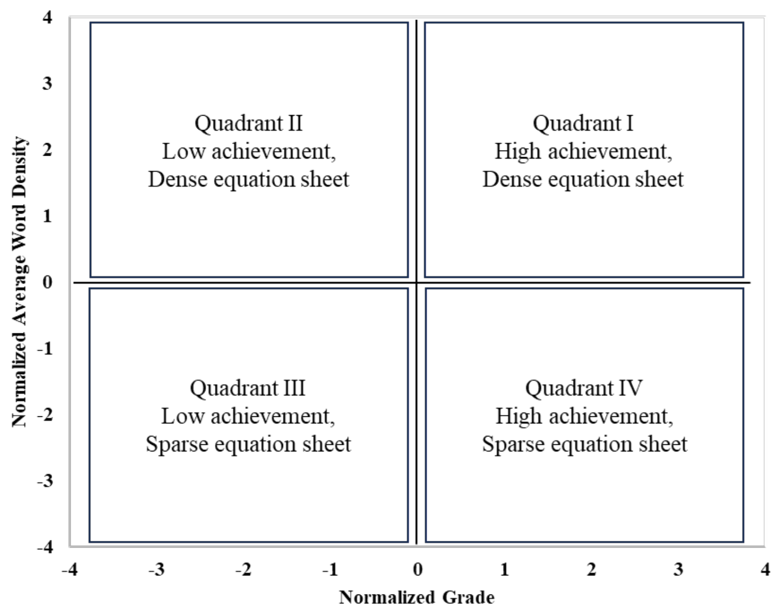


Figure 1. Quadrants showing relationship between normalized grade and average word density of the equation sheet.

sparse equation sheet in Quadrant 3, and a high grade with a sparse equation sheet in Quadrant 4 (Figure 1). Each of these quadrants have connotations related to student effort and performance, with Quadrant 4 indicating the highest level of internalization of the material (i.e., high relative achievement on the graded event with the lowest relative amount of information on the equation sheet).

In addition to this classification, specific data points were judged to be significant for falling outside of one standard deviation from the mean (Equation 1) on the vertical axis (Figure 2). This significance was selected empirically to focus the importance on the average word density for the equation sheets and to filter out performances that were similar across the study population. This choice is not meant to imply that the grades for the students in the population are not important. Rather, it is meant to recognize that the performance of a student population on a major graded event will regularly have outlying low grades (for a range of reasons), thus the focus on the average word density.

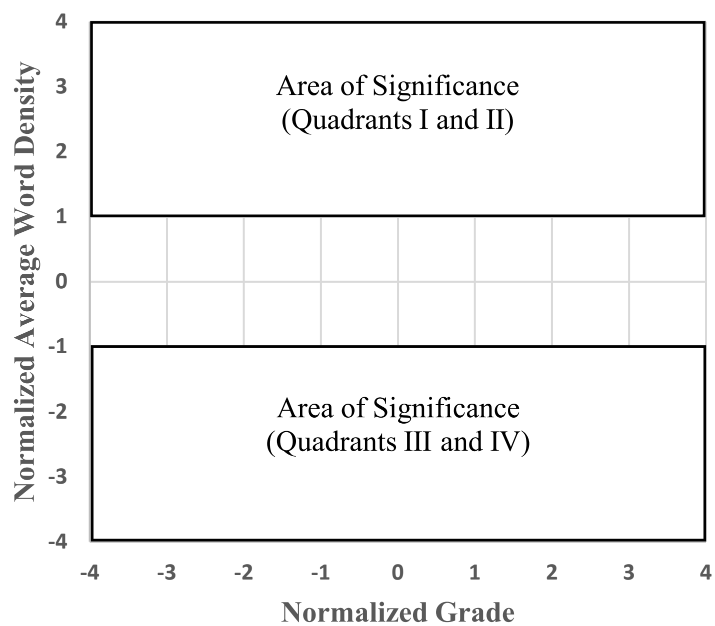


Figure 2.

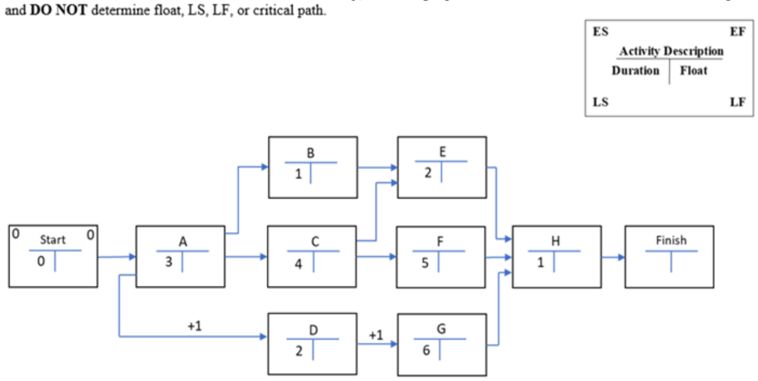
Experiment and Data Set

This research uses the final exam grade in *CE450: Construction Management* for comparison against the average word density on the student-generated equation sheet that each student was allowed to bring into the final exam. This particular course was chosen for this study based on its consistent allowance of student-generated equations sheets. The exam consisted of the typical questions on the fundamentals of construction management (Figure 3) for such topics as project delivery methods, estimating, network diagrams, Gantt charts, resource leveling, and earned value analysis. Students were allotted three and a half hours to complete the exam and could have one 8 ½ inch x 11-inch equation sheet (front and back sides available). The equation sheet had to be filled out manually and students were not allowed to print or share equation sheets.

(5 points) You are the contractor on a small construction project. Liquidated damages are set in the contract at $2,000 per day. You realize that you are going to finish the project 10 days late. If you crash the project, you can finish on time. Of the options below, what is the highest cost where it would still be worthwhile to crash?

a. $16,000
b. $18,000
c. $20,000
d. $22,000

2. (20 points) Given the durations (days) of each activity in the network diagram (not the same as problem [1]), conduct a **FORWARD PASS**. Determine **ES** and **EF** of each activity, and **total project duration**. **DO NOT** conduct a backwards pass and **DO NOT** determine float, LS, LF, or critical path.



Total Project Duration (days):

2. (20 points) Given a resource constraint of **8 laborers**, generate a resource leveled Gantt chart ***without delaying the project AND without increasing the number of critical tasks*** in the project. Use the method outlined in the Seabee crew leader handbook (i.e., technique taught in class) to annotate your adjusted schedule. Note that this Unconstrained Gantt chart is different from the one in problem [1].

Max Labor Size 8

ID	Duration (Days)	Precedence	Resource	Float	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
A	3	xxx	3	0	3	3	3																			
B	3	A	4	6				4	4	4																
C	5	A	2	3				2	2	2	2	2														
D	5	A	4	0				4	4	4	4	4														
E	4	C FS+2, D FS+2	3	5											3	3	3	3								
F	6	D FS+3	5	2												5	5	5	5	5	5					
G	4	D FS+4	2	0													2	2	2	2						
H	2	B FS+7, C FS+8, E, F, G SS+7	3	0																				3	3	
			Unconstrained Labor		3	3	3	10	10	10	6	6	0	0	3	8	10	10	7	7	5	0	0	3	3	
			Leveled Labor																							

Figure 3. Sample problems similar to the problems on the final exam.

The data set for this research included 210 students that took the course over two semesters between the fall of 2022 and the spring of 2023. Of note, the same instructor managed the course and the development of graded events for both semesters, and the same final exam was used for each semester. However, multiple instructors taught various sections (groups of students) in each semester and contributed to grading the final exam. As a result, it was not possible to have a single instructor grade the entirety of the final exam.

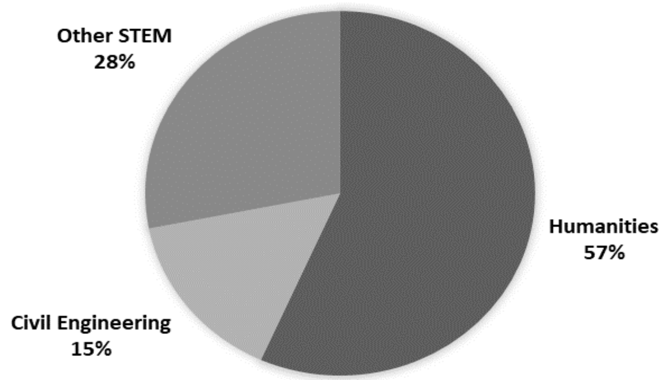


Figure 4. Population percentages based on academic major.

A unique aspect of this course (along with all other courses in the various three-course engineering sequences) is that many non-engineering students are enrolled. The diversity of academic backgrounds results in over half of the students coming from humanities-based disciplines (Figure 4). The civil engineering enrollees are separated from the other STEM-based majors due to the course falling within the civil engineering program. Figure 5 shows the final exam grades in the course. While it shows a degree of skewness, the distribution indicates that the students largely performed well on the exam.

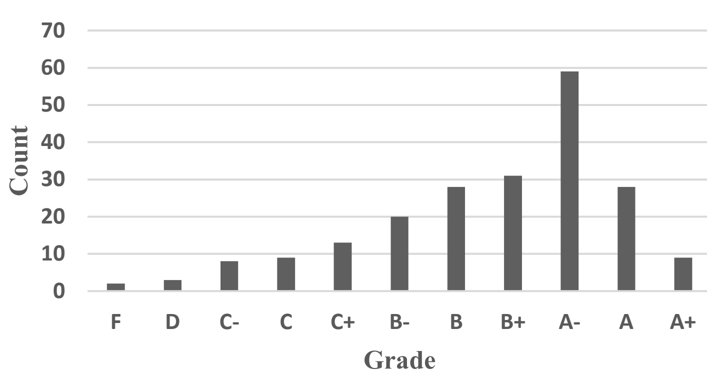


Figure 5. Distribution of final exam grades in CE450.

The average word density was collected manually by evaluating the number of words in six separate 2-inch x 2-inch samples of each equation sheet. Three samples were taken on each side of the equation sheet before averaging the values, with the first sample taken at the top-left of the page, the second in the middle, and the third at the bottom-right of the page. This sampling approach was used to account for the recording pattern of the students, as it would capture variance in the recording habits of the students in cases where the entire page was not used. Additionally, all words in the sampling space were counted along with partial words where the researcher was able to delineate the given word. Also, elements of formulas were counted as words, along with labels and axis values of graphs (but graphs themselves did not count). This approach was selected as the most efficient, consistent, and realistic way to record the average word density, although the authors recognize that it is not as ideal as tallying the entirety of the information on the equation sheet. The same researcher conducted the evaluation for each of the 210 equation sheets. The distribution of the average word density values (Figure 6) shows a noticeable skewness with a modal value around four words per square inch (wpi), a maximum value around 18 wpi, and a minimum value of less than 1 wpi.

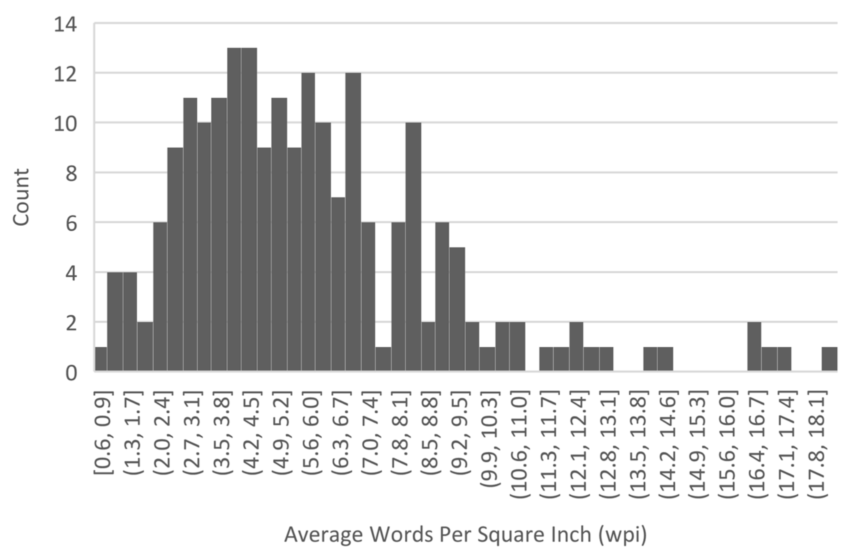


Figure 6. Distribution of average word density

Results

The raw data for the relationship between average word density with the associated grade on the final exam are shown in the following graph (Figure 7). Each population group is indicated in the figure with a different symbol, and each axis is included along the mean value (normalized value of zero) to better identify which of the four quadrants each data point falls into. Proportions of the study population in each quadrant are also included (Table 2).

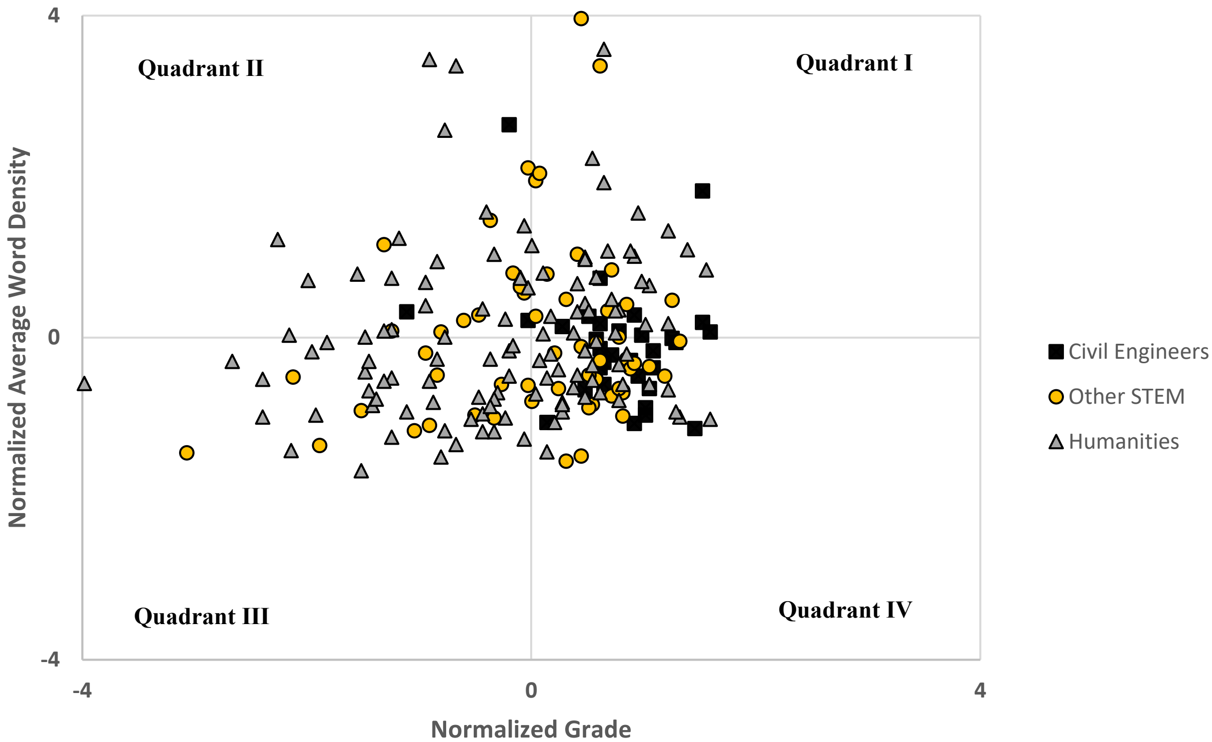


Figure 7. Relationships between grade and average word density for this data set. All values are normalized.

Table 2. Distribution of the study population in each quadrant.

Quadrant	Description	Proportion
I	High grade, High WPI	25%
II	Low grade, High WPI	18%
III	Low grade, Low WPI	24%
IV	High grade, Low WPI	33%

The charts in Figure 8 highlight the data in each population group that falls outside of one standard deviation with respect to either the normalized grade or the average word density of the student-generated equation sheet (indicated by the shaded regions). While one standard

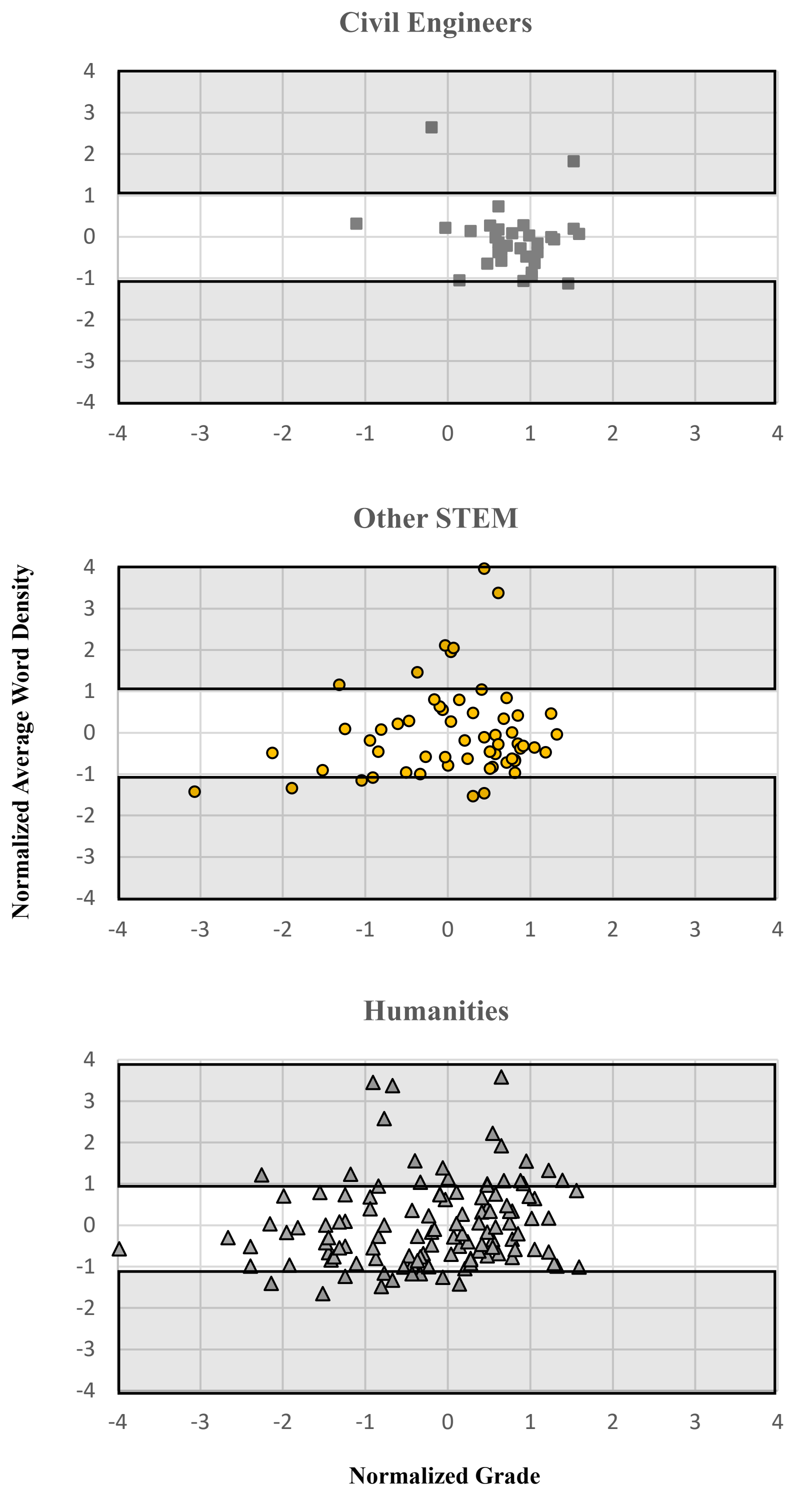


Figure 8. Areas of concern are highlighted on the graphs.

deviation may seem somewhat arbitrary, the authors chose this value empirically to filter out performances that were common across the study population.

Discussion

The desired outcome for this data set (and future data sets) is that the students fall into Quadrant IV. This outcome would indicate that they are generally internalizing the course material and are able to demonstrate proficiency without significant references. This relationship is akin to anecdotal stories from educators that high performing students having sparsely populated self-generated equation sheets. It is promising to see, in this case, that a third of the population in this study falls into this category (Table 2).

It is worth noting, however, that the overall course performance is high, as indicated by the peak of the distribution of raw grades in the course occurring at the A- level (Figure 5). There is clearly variability (which cannot be ignored) in the difficulty of this course compared to others. Nevertheless, it is evident that the data shows a healthy percentage of the study population reflecting a strong performance with a lower-than-average word density on their equation sheets. This trend is promising given that the WGCD statements (Table 1) for this course focus on application rather than simple knowledge retention, for example.

Of the students that performed above the average with respect to the final exam grade, a quarter of the study population is grouped into Quadrant I. This shows that many students, despite being generally proficient with the course material, created self-generated equation sheets with a higher-than-average word density. A qualitative analysis of these equation sheets revealed an inclusion of a significant number of definitions and explanations of key terms. While this does not initially seem problematic, it may indicate that these students expected to be evaluated based on knowledge and recall as opposed to analysis and application of principles. The authors recognize that there is nothing inherently wrong with students including definitions on their equation sheets, but it may indicate that the lesson activities and assessment strategy in the course prior to the exam could be more clearly communicated.

While some of the outlying data points skew the data such that less than half of the population falls below the average grade on the event, this portion of population more heavily favors Quadrant III compared to Quadrant II. The population in Quadrant III, with a lower-than-average grade and a sparsely filled equation sheet, unfortunately connotes at best a struggle with the course material and at worst a lack of effort. In contrast, the population in Quadrant II shows a poor performance albeit with effort placed into the equation sheet. While the smallest cohort of the data set, it still potentially reflects a similar characteristic to those students in Quadrant I by filling the respective equation sheets with definitions and facts.

Overall, the relationship between average word density and exam performance (Figure 7) struggles to offer a deeper insight than an analysis of the grade distribution for the event (Figure 5). Given the skewness of the grades, around 58% of the students performed higher than the average score on the exam. While it is recognized that other measures, such as mean or mode, could provide a different perspective on central tendency of the data, it is interesting that this same mean is maintained when evaluating the exam performance given the classification of the average word density. In other words, an evaluation of all students with a higher-than-average word density (Quadrants I and II) shows that 58% of the students scored higher than average on the exam. Likewise, the collection of students in Quadrants III and IV have nearly the same relationship with their exam performance. This relationship discredits the notion that the density of material on equation sheets can adequately indicate exam performance.

Lastly, observations of concerning data (outside of one standard deviation for average word density) show higher occurrences for students with academic majors from external departments (Figure 8). The vertical spread of the population groups from external departments is generally two standard deviations (in total) higher than those from the civil engineering population group, and it tends to favor the area of significance related to high average word densities. A potential explanation is that the civil engineering students have higher familiarity with the teaching pedagogy and approach to evaluation and assessment from the instructors. This familiarity results from the civil engineer population typically taking the course later in their careers after exposure to many other courses from the department. This exposure is something that, by the nature of their programs of study, the populations from other STEM and humanities departments do not have. Of note, the grades on this event did not account for overall academic performance (such as cumulative grade point averages) of the students, as it would be outside the scope of this work. However, such an analysis could potentially provide explanations for the disposition of the data in these population groups.

Conclusion

Given the results of this work, the authors return to the original research questions:

1. To what degree is there a relationship between the amount of content that a student places on a self-generated equation sheet and the performance on the associated graded event?
2. Do the relationships between the equation sheet and graded event reveal anything about our approach to teaching and evaluation?

To answer the first question, the data shows that the relationship between equation sheet word density does not offer insight into exam performance beyond the raw distribution of grades on the graded event. Ultimately it does not fit the age-old belief that minimal equation sheet information correlates to high grades with crammed information correlating to low grades. The results from this study show a full range of performances on the exam regardless of the density of information on the equation sheets.

This does not mean, however, that there is no value in this information. With an overall goal to get our students to internalize the information we present during the progress of an academic course, the knowledge of the proportion of students with high grades and low word density equation sheets can serve as a benchmark. It is one of a number of quantifiable assessment tools that can arguably indicate a degree of mastery of the course material. Further, it can provide a disposition of students regarding the demonstrated effort in the course and comfort level with being evaluated on the course material.

Regarding the second question, the differences in the results between the population groups in the study shows that, as instructors, we must pay careful attention to the diversity of academic backgrounds in our classrooms. As was the case with this data set, students with academic backgrounds that were different from the academic discipline of the course had the largest instances (proportionally) of extreme values (both grades and word density). These students require us to exercise careful thought about delivering material in a suitable way to the varied learning types in the seats of our classrooms.

Sadly, this study also highlights that portions of the students in our classrooms do not put forth effort to master the information that we teach. While it is incumbent upon the instructor to present the material in the best way possible, there is still a degree of individual responsibility on behalf of the students to achieve a level of performance in line with their individual potential. This does not mean that we disregard those students that do not appear to put forth effort. Rather, we owe it to them to do our best to provide purpose and motivation to the best of our abilities.

Works Cited

[1] Office of the Dean, Academic Program: Curriculum and Course Descriptions, West Point, New York: United States Military Academy, 2023.

[2] Office of the Dean, Dean's Policy and Operating Memorandum (DPOM) 05-07: Assessing and Improving Student Learning in the Academic Program, West Point, New York: United States Military Academy, 2019.

[3] D. M. Settlage and J. R. Wollscheid, "An Analysis of the Effect of Student Prepared Notecards on Exam Performance," *College Teaching,* vol. 67, no. 1, pp. 15-22, 2019.

[4] K. H. Larwin, J. Gorman and D. A. Larwin, "Assessing the Impact of Testing Aids on Post-Secondary Student Performance: A Meta-Analytic Investigation," *Educational Psychology Review,* vol. 25, pp. 429-443, 2013.

[5] K. L. Dickson and J. J. Bauer, "Do Students Learn Course Material During Crib Sheet Construction?," *Teaching of Psychology,* vol. 35, pp. 117-120, 2008.

[6] Y. Song and D. Thuente, "A Quantitative Case Study in Engineering of the Efficacy of Quality Cheat-Sheets," in *IEEE Frontiers in Education Conference*, El Paso, TX, 2015.

[7] A. Gharib, W. Phillips and N. Mathew, "Cheat Sheet or Open-Book? A Comparison of the Effects of Exam Types on Performance, Retention, and Anxiety," *Psychology Research,* vol. 2, no. 8, pp. 469-478, 2012.

[8] M. J. Sanborn, K. T. Purchase and B. E. Barry, "Kicking Out the Crutch: The Impact of Formula Sheets on Student Performance and Learning," in *Annual Conference of the American Society of Engineering Education*, San Antonio, TX, 2012.

[9] J. K. Vessey and W. Woodbury, "Crib Sheets: Use with Caution," *Teaching Professor,* vol. 6, no. 7, pp. 6-7, 1992.

[10] S. A. Danielian and N. T. Buswell, "Do Support Sheets Actually Support Students? A Content Analysis of Student Support Sheets for Exams," in *Pacific Southwest Section Meeting of the American Society for Engineering Education*, Los Angeles, CA, 2019.

[11] J. A. Shaw and M. G. Almeida, "Does the Quality of Student Crib Cards Influence Anatomy and Physiology Exam Performance?," *HAPS Educator,* vol. 22, no. 1, pp. 50-54, 2018.

[12] S. Hamouda and C. A. Shaffer, "Crib Sheets and Exam Performance in a Data Structures Course," *Computer Science Education,* vol. 26, no. 1, pp. 1-26, 2016.

[13] C. Cannonier and K. Smith, "Do Crib Sheets Improve Student Performance on Tests? Evidence from Principles of Economics," *International Review of Economics Education,* vol. 30, pp. 1-13, 2019.