Choice of Major and Career Aspirations of First-Year ECE Students

Dr. J.W. Bruce, Tennessee Technological University

J.W. Bruce is with the Department of Electrical & Computer Engineering at Tennessee Technological University in Cookeville, Tennessee USA

Dr. Mahnas Jean Mohammadi-Aragh, Mississippi State University

Dr. Jean Mohammadi-Aragh is an assistant professor in the Department of Electrical and Computer Engineering at Mississippi State University. Dr. Mohammadi-Aragh investigates the use of digital systems to measure and support engineering education, specifically through learning analytics and the pedagogical uses of digital systems. She also investigates fundamental questions critical to improving undergraduate engineering degree pathways. . She earned her Ph.D. in Engineering Education from Virginia Tech. In 2013, Dr. Mohammadi-Aragh was honored as a promising new engineering education researcher when she was selected as an ASEE Educational Research and Methods Division Apprentice Faculty.

Choice of Major and Career Aspirations of First-Year ECE Students

INTRODUCTION

Typically, university engineering study is categorized into specialty areas, e.g. civil, chemical, computer, electrical, mechanical, etc. Engineering students are asked to select a major in one of the engineering specialty areas upon matriculation, or soon thereafter. Previous research has shown that significant factors influencing choice of major for college students include (1) general interest subject; (2) family and peer influence; (3) assumptions about introductory courses, (4) potential job characteristics, and (5) characteristics of the major. The student's decision on choice of major is often difficult because traditional university-aged students have little to no direct experience with the engineering profession or practicing engineers. Some universities confront this problem with a common first-year engineering experience, wherein engineering majors are given the opportunity to explore the specialty areas and make a more informed decision. Other institutions, including the authors', implement discipline-specific first-year experience to allow students to learn specialty skills, and more immediately identify with their chosen profession. Regardless of the approach, the desire is that students find their professional path quickly to avoid delays in graduation and increased student debt.

The authors teach an introductory discipline-specific course in electrical and computer engineering, in which most students have declared their intended academic degree program. A small number of students enrolled have not declared their desired engineering major, or are currently classified as some non-engineering major. While the course used in this study is a freshmen/first-year introductory course, the course is required of all electrical and computer engineering graduates. As a result, the course enrolls students classified as freshmen, sophomore, juniors, and seniors. The authors collected data on student career aspirations from almost 600 students over a four-year period with a question that demanded an open-ended, free-form prose response. Students answered in their own words. The student responses have been analyzed with textual data mining techniques and several sentiment analysis algorithms to ascertain common ideas and the sentiment of the student responses. These results were analyzed with populations of the study group based on their declared major and university classification.

BACKGROUND

ABET defines engineering as “*the profession in which knowledge of the mathematical and natural sciences gained by study, experience, and practice is applied with judgment to develop ways to utilize, economically, the materials and forces of nature for the benefit of mankind*”. With this definition, one can view the different disciplines of engineering as bringing to bear mathematics and their respective natural sciences to form a solution. Chemical engineers would employ lots of chemistry, civil and mechanical engineers would naturally use lots of Newtonian physics, electrical engineers would employ solid-state physics and electro-magnetics, computer engineers would use more computer science, etc. With such an interpretation, engineers in the different sub-disciplines are very much alike with some small differences. It would be reasonable to assume that students entering study of these fields of engineering would also be very similar with some small differences. Engineering students are a population are not completely homogeneous. Research bears this out.

In [Pot2013], the authors found students enrolled in specific engineering disciplines expressed different affinities for different fields of science, and were varied in their perceived practicality of the different engineering disciplines. For electrical and computer engineering (ECE) students, the authors found that the typical student preferred physics slightly more than most other engineering fields, and reported a self-perceived lower skill level but greater interest in mathematics. ECE students also report a very high interest in “inventing/designing things”, and viewed the work of electrical and computer engineers as being broadly/globally applicable.

A large study [Bei2016] noted that open-ended responses from engineering students indicated that 16% of students reported that “helping people” was a factor in their decision to study engineering, and 7% of respondents reported that “helping people” was the primary or sole reason for their choice. Indeed, these university students agree with impact of engineering as expressed in ABET’s definition of engineering. Students studying biomedical, environmental, materials, and civil engineering were more likely to so strongly driven with altruistic motives, and electrical, computer and aerospace engineering being less empathetic. The study in [Bei2016] also examined the degree to which students perceived certain engineering disciplines help people/society. Students reported that that engineering disciplines that prioritize helping other most are chemical and biological engineering. Civil and environmental engineering place a moderate priority on helping others, while electrical, computer, and mechanical engineering have the lowest priority for creating solutions to humanitarian problems. It seems that we have a ways to go in educating the the public and potential students about how all fields of engineering strive to improve the lot of all humankind.

Another study [Ryn1988] found that both engineering student and working engineers aspire to ultimately serve in management or leadership roles. (The study did not differentiate between the different fields of engineering.) Such a conclusion would likely not surprise an engineering educator as most engineering students are intelligent, highly motivated, and exhibit good leadership ability, even as young adults. The study, perhaps, more surprisingly found that student engineers aspire to leadership and management at a greater percentage than working engineers.

As to career aspirations and persistence, a number of studies [Cav2007] [Ben2015] [Pal2010] [Chr2014][Alp2008][Shi2003] found that students’ abilities, perception of abilities, especially in mathematics play a big part in persistence. Another large contributing factor is the students’ aspirations and how well the discipline – or more accurately, their perception of the discipline – matches with their career and personal aspirations. To improve retention, engineering programs need to ensure that students recognize how their aspirations and interests align with their chosen field early in their studies. Toward this end, an accurate picture of student aspirations is needed.

Sentiment Analysis

Sentiment Analysis is a sub-field of natural language processing. In short, it is data mining of free-form textual data arising from human communications. The aim is to identify and extract opinions contained within text data. Sentiment analysis desires to quantify attitude, sentiment, and emotion of a speaker or writer based on a computational treatment of text. Mostly often, the textual data is unstructured making automated analysis more difficult. A well-known use of sentiment analysis is by motion picture studios monitoring the social media "buzz" about their latest movie release. Social media such as Twitter or Facebook are mined for references to a studio's latest movie titles, and the sentiment or emotion expressed in proximity to the title are analyzed. The goal is to determine if the public is positive or negative in their regard for the film. This information can provide early indicators to a film's potential market success, and advertising strategies can be altered based on the results. A particular challenge for textual data mining and sentiment analysis from social media data is the unstructured nature and the huge volume of data to be analyzed.

Though it may seem easy on first blush, sentiment analysis is quite difficult. Understanding emotion expressed in text by a computer is not easy. Sometimes, even humans can get misled, even though people have a lifetime of experience with human communications. A clear example of human communication misunderstanding would be emails that you have read and completely misinterpreted the intent or emotion of the author. Text may contain multiple conflicted sentiments all at once. For instance, "The candidate's speech had lots of great ideas, but we still can't afford any of them." The above sentence contains sentiment with opposing polarities: positive ("lots of great ideas") and negative ("can't afford any of them"). Should this sentence, as a whole, be construed as positive or negative? This is a key challenge in sentiment analysis.

Computers methods can struggle to ascertain emotion or sentiment, especially with figurative speech or idioms. We, as humans, use idioms all of the time to convey a very specific emotion or feeling, although the words themselves actually mean something completely different. Consider the statement "all I will say is that life with Charlie is interesting." Here, the word "interesting" could be extremely positive or extremely negative, or almost any degree in between. Context, prior statements, vocal tone (if spoken), and body language convey much more information that humans readily understand, but computers do not. For example, you might hear someone exclaim "that's wicked!" when describing a friend's new smart phone. The word "wicked" would generally be viewed as a negative adjective meaning "evil" or "cruel". In this context, the exclamation is most likely the speaker expressing a strong sense of approval or admiration. Much effort has been exerted to improve textual data mining for sentiment analysis to help computers overcome such challenges.

A common way to determine sentiment of a body of text is to calculate the sum of the sentiment of the text's individual words. That is, sentiment analysis is often based on its unigrams, i.e., single words. A sentiment analysis lexicon will contain many English words with each word assigned a numerical score representative of the word's sentiment. Most lexicons contain data for determining the "valence" emotive sentiment. The valence of a word, feeling, idea or situation describes the intrinsic "badness" or "goodness" or any degree in between. Bing Liu analysis

[Liu2010] is one straight-forward sentiment analysis. The Bing Liu lexicon categorizes words in a binary fashion. Every word in the Bing Lui is either “positive” or “negative”. While simple, Bing Lui sentiment analysis will not account for varying degrees of emotive response of words or ideas. For example, “tolerate”, “like”, “love” and “adore” all express your differing positivism toward an idea or object. Bing Lui would rate all four as being equal. Other sentiment analysis approaches allow for varying degrees of emotive response. The techniques are the most commonly used today. Furthermore, sentiment analysis lexicons may also contain sentiment values along other emotive “dimensions”, such as joy, anger, sadness, arousal, and so forth.

A popular and high-performing sentiment analysis method is “Valence Aware Dictionary and sEntiment Reasoner” (VADER). VADER is a lexicon and rule-based sentiment analysis tool that was specifically developed for assessing sentiment expressed in social media [Hut2014]. VADER assigns each lexical feature (word) a classification, where each word is labeled as being varying degrees of positive or negative. VADER also accounts for punctuation modifiers. Exclamation points increase the positivity or negativity of the associated words. Words that expressed in all caps – the modern internet notation for shouting – are also assumed to increase the degree of sentiment. VADER further adjusts sentiment degree with modifier adverbs like “very”, “extremely”, “sort of” etc. Finally, VADER will reverse sentiment direction in the presence of words like “not” or “but”. For example, “The food was not delicious.” “Delicious” would be considered positive, but the modifier “not” flips the sentiment to negative. The VADER sentiment analysis algorithm takes such negation into account. VADER has several advantages over other methods of sentiment analysis: VADER is fast; it is not very computational intensive; and VADER doesn’t require any training data because it is based on its included sentiment lexicon. A drawback of the published VADER sentiment method is that its is targeted to quickly decide the sentiment of very limited amounts of text in social media – think: tweets, Facebook posts, and Instagram captions. In other words, VADER strives to judge overall sentiment in the first few dozens of words. Analysis of sentiment by VADER of longer texts tends to saturate at the extremes: very positive or very negative.

Another commonly used sentiment analysis was proposed by Finn Nielsen [Nie2011]. This approach, commonly called AFINN, uses a lexicon of English words rated for valence with an integer scale between minus five (negative) and plus five (positive). Steven Loria has spearheaded the open-source community development of another sentiment analysis library for natural language processing called TextBlob [Lor2019]. TextBlob does part-of-speech tagging, noun phrase extraction, sentiment analysis, and classification. For sentiment analysis, TextBlob considers sentiment over the range of -1 (negative) to +1 (positive). Furthermore, TextBlob also quantifies the perceived subjectivity of words over the range of 0 (objective) to +1 (subjective). Both AFINN and TextBlob use their lexicons to determine the overall sentiment of a text sample by averaging the sentiment of individual words. Neither approach tries to account for degree modifying words or negation.

Going beyond simple valance, Bradley and Lang created a normative emotional rating English words in their ANEW 2017 lexicon [Bra2017]. Inspired by existing emotive ratings for pictures, sounds, and emoticons, they had a large number of “Introductory Psychology” students, where the study group was gender-balanced, rate a lexicon of thousands of English words according to three emotive dimensions: valance, arousal, and dominance. Subjects rated each word from one to nine for each emotive dimension. A word's "valence" ranges from unpleasant (1) to pleasant (9). "Arousal" quantifies a word's affect on the subjects' excitability from calm (1) to excited (9). "Dominance" measures the emotive response to a word's sentiment of control. Dominance ranges from feelings characterized as completely controlled, cared-for, awed, submissive, or guided (1) to completely in control, influential, important, dominant, and autonomous (9). Bradley and Lang claim that valance and arousal responses to most words is strong, while the dominance emotional response is "less-strongly related". ANEW2017 does not specify how the overall sentiment is calculated. ANEW2017 is simply a lexicon for use with other algorithms. Typically, the ANEW2017 lexicon is used in a simple average methods like AFINN and TextBlob.

APPROACH

The authors teach an introductory course in electrical and computer engineering which was created to specifically address (1) provide an orientation and early success skills for university life, (2) introduce ethical considerations in engineering, (3) introduce the profession of engineering, and specifically, electrical engineering and computer engineering, and (4) give early technical and hands-on skills required of EE and CmpE majors. Students in the course have predominantly already selected computer engineering or electrical engineering as their field of study; however, a number of students enrolled in the first-year course are exploring the fields of computer and electrical engineering in their search for a major. As the introductory course is a prerequisite to later ECE courses, it is taken very early in the student's university tenure. Freshman take the course in their first or second semester at the university. Transfer students almost always take the course in their first semester at the institution because the course is prerequisite to following courses which compose the longest prerequisite chain through the program.

In the second lecture period of the course, the authors present a question to the class: "What would you do on a daily basis in your dream job after your graduate from electrical or computer engineering?" Students respond in free-from prose on paper that is collected at the end of the period. Students can respond in any manner, with any words of their choosing. The first meeting of the course (prior to posing the question above), student experience a series of guest speakers about department policy on advising, computer usage, building access, etc. Nothing of significance about the profession of engineering, electrical engineering, or computer engineering has been discussed before the question is posed. When queried in the fall semester offerings, nearly every student in the course is in their first week as a new student at the university. The spring semester cohort is predominantly in their second semester at the institution, and, in general, have not taken any other engineering courses. It is posited that students' aspirations and impressions of the engineering profession have not be appreciably influenced. The intent was to collect data from the "unaltered" new electrical or computer engineering university student.

Data was collected in this way from 577 students over a four-year period. As the course described here applies toward graduation only for electrical engineering (EE) and computer engineering (CmpE) majors, a vast majority of the study population were majoring in EE or CmpE. Table 1 shows the breakdown of the study population by declared major as reported by the university data management system upon the first day of the course. Engineering-Undeclared are students enrolled in the university's College of Engineering, but have not selected a specific program of study yet, or enrolled as a "pre-engineering" student in the College of Engineering. "Pre-engineering" students often are classified as such because they are remediating some foundational area such as mathematics, English, etc. Undeclared students have not selected or identified any particular major to the university.

Since the course only provided graduation credit for EE and CmpE degrees, it is assumed that students enrolled in the course who are not EE or CmpE majors fall into one of three categories:

(1) transferring into ECE and enrolled in the course *before* they officially changed majors,

(2) already enrolled in the course but changing out of EE/CmpE with their major change already having been processed, or

(3) taking the course because of curiosity in the ECE profession with no current (or immediate) plans to change their major to EE or CompE.

The membership of non-ECE students into these three categories was not ascertained.

N=577	***Number***	***Percentage***
Computer Engineering	219	38%
Electrical Engineering	277	48%
Aerospace Engineering	2	0.3%
Chemical Engineering	1	0.2%
Computer Science	2	0.3%
Engineering-Undeclared	8	1.4%
Mechanical Engineering	3	0.5%
Software Engineering	5	0.9%
Business Administration	1	0.2%
Geoscience	1	0.2%
Physics	1	0.2%
Undeclared	49	8.5%
Could not be determined	8	1.4%

Table 1: Declared major of student population

Since all students in the both the EE and CompE programs must ultimately take the course described here, it is common to see students in a variety of places along their academic career in the course. Table 2 shows the class standing of the students in the course. Data reported here was student's classification on the first day of the course. The data available does not allow us to easily determine each student's true higher education background. For the purposes of this study, it is *assumed* that students enrolled in the course classified as freshmen are "true freshman" – the authors' institution is the first (and only) institution of higher education in which they have enrolled. This is likely correct for a vast majority of the students in the study, as advising steers them into the introductory course in the first or second semester of study. Non-freshman are very likely to be transfer students from community college as approximately 40% of all EE and CmpE majors in the program transfer to the university from community college. It is possible that some of the non-freshmen students simply delayed taking the course beyond the norm, or that these non-freshmen students have transferred into EE or CmpE major after some semesters of study in another major. The authors are also aware of a few first semester university students who are sophomore and juniors due to AP course credits and/or dual-enrollment during their high school years. Identifying these fairly rare cases is very difficult with current university data systems.

N=577	***Number***	***Percentage***
Freshman	221	38.3%
Sophomore	162	28.1%
Junior	154	26.7%
Senior	32	5.5%
Could not be determined	8	1.4%

Table 2: University classification of student population

The student provided responses to the question of their aspirational career duties were electronically captured and analyzed. Some of the student hand-written responses were illegible and omitted. Analysis of the student responses followed fairly standard form for textual data analysis. First, "stop words" were removed. In natural language processing, useless words (data) are called stop words. Stop words are common words typically omitted in search engines and often ignored with alphabetizing titles. Common stop words are "the", "a", "an", "in", "on", etc. The next step of analysis was to process the responses for parts-of-speech. Each unigam was examined and marked as to its part of speech: "noun", "verb", adjective", "adverb", etc. Word clouds showing the relative frequency of words were made. Texts analyzed included full student responses, minus stop words, as well as nouns only, and verbs only. Finally, full text minus stop words were analyzed for overall sentiment using VADER, AFININ, TextBlob, and ANEW2017 lexicons for students by majors and classification.

ANALYSIS & DISCUSSION

In order to see if particular ideas or objects are more common among several different populations, a word cloud was generated over the nouns and verbs responses. A word cloud is a common pictorial representation of textual data where a higher frequency of a particular word causes it to be rendered larger. In short, small words in a word cloud have a low number of occurrences, and larger words have higher numbers of occurrences in the associated text. Figure 1 and Figure 2 show word clouds generated from nouns used in response by the study population whose classification are ECE (either EE or CmpE) and non-ECE (everything else), respectively.

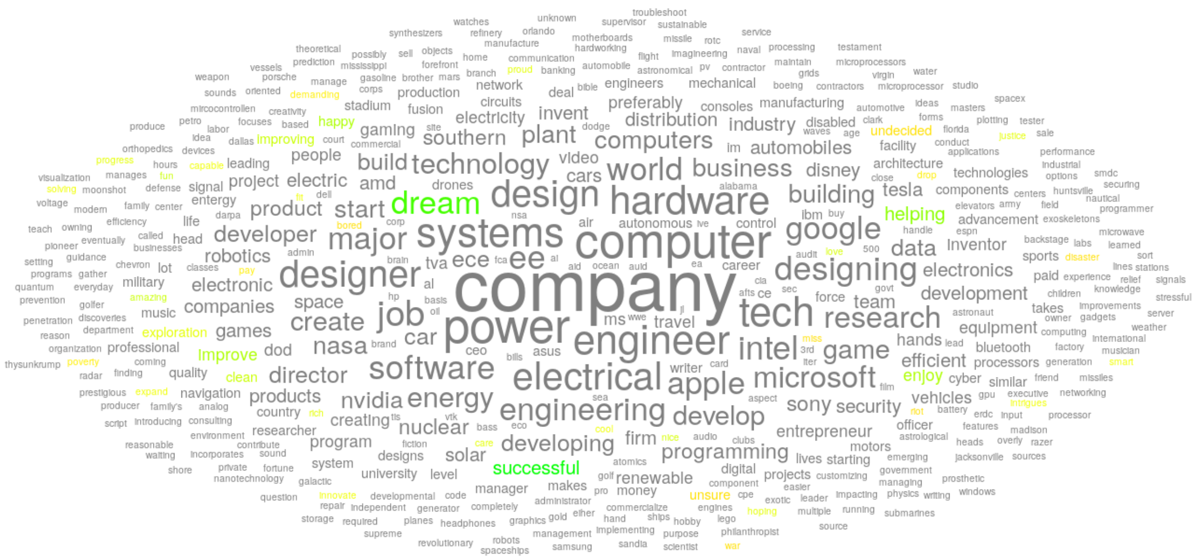

Figure 1: Nouns word cloud from responses of electrical (EE) and computer (CmpE) engineering majors

Figure 2: Nouns word cloud from responses of students whose declared major is not electrical (EE) or computer (CmpE) engineering

Comparing Figures 1 and 2, several high-frequency nouns are common as to both as one might expect in response to the question posed:

Engineer, company, design/designer, job, computer, hardware, tech

Some differences that can be easily noted between Figures 1 and 2. ECE majors (Figure 1) mention the nouns *systems, power, energy, Google, Apple, Microsoft, research, world* in greater proportions than non-ECE majors (Figure 2). ECE majors – with higher frequency – describe their aspirations in terms of specific corporate entities, and mention *energy, power*, and *research* – all very broad concepts common to all aspects of science and engineering. Non-ECE majors (Figure 2) mention the nouns *gaming, software, vehicles, job, engineer* in greater proportions than ECE majors (Figure 1). Non-ECE majors tended to describe their dream job with nouns that are technical applications (gaming, software, vehicles) typically viewed as being more tangentially related to electrical and computer engineering. Also, it should be noted that non-ECE also were more likely to write about *job* and being an *engineer* as opposed to specific aspects of electrical and computing technology and disciplines. This last observation tends to agree with [Bei2016] that found that electrical and computer engineering students are less likely to report an emphasis on the profession on helping people. Interestingly, the words clouds do not give much indication of aspirations of leadership and management as found in [Ryn1988].

Figure 3 and Figure 4 show words clouds generated from nouns used in response by the study population whose classification are freshman and non-freshman, respectively.

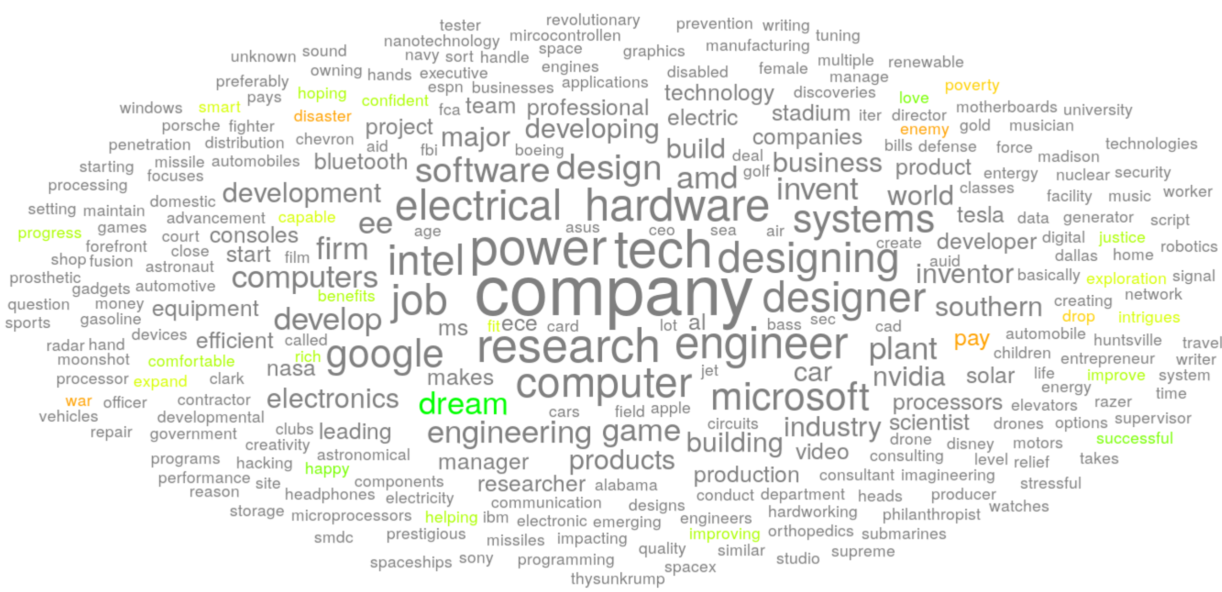

Figure 3: Nouns wordcloud from responses of freshman

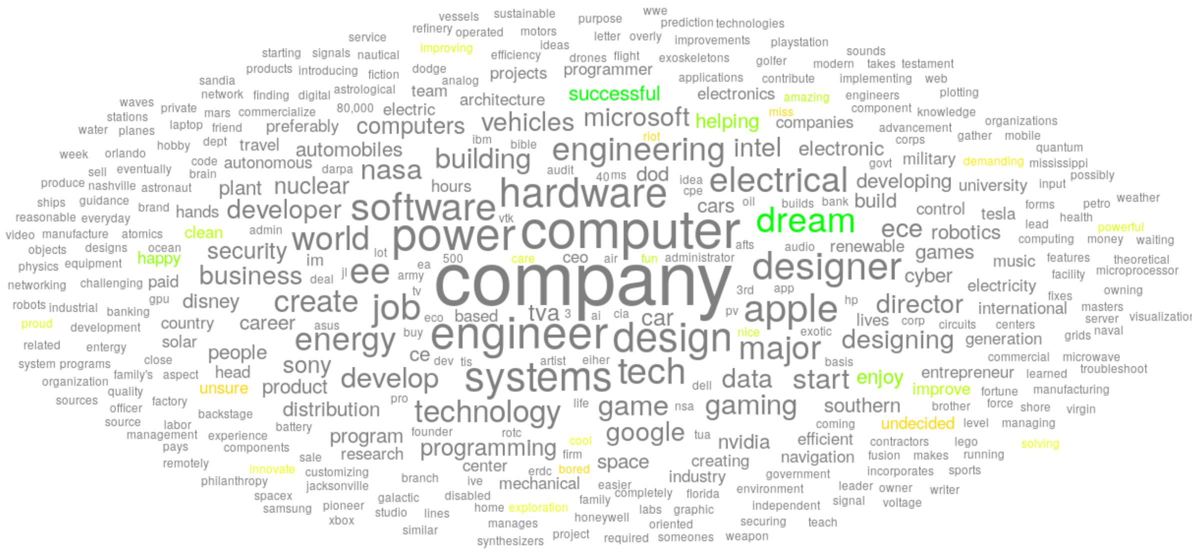

Figure 4: Nouns wordcloud from responses of non-freshman (sophomores, juniors, or seniors)

Examining Figures 3 and 4 shows that differences between freshmen and non-freshmen are subtle at best. Both populations frequently mention

company, job, engineer/-ing, design/-er, systems, power, hardware, software, tech, business

Few significant differences are apparent in the noun responses. The largest difference among the most common nouns is that freshmen (Figure 3) cited *research* as an aspiration in their future career more often than non-freshman (Figure 4).

Sentiment Analysis

The textual data responses were analyzed using VADER, AFFIN, TextBlob, and ANEW2017 methods and lexicons for sentiment for comparison. The entirety (all parts of speech) of the students' responses minus stop words were input to the sentiment analysis algorithms. VADER analysis rates words as being positive, negative, or neutral. Table 3 show the percentages of words of each VADER sentiment in the responses by university classification.

	Positive	***Neutral***	***Negative***
Freshmen	15 %	82 %	3 %
Sophomore	7 %	90 %	2 %
Junior	13 %	85 %	2 %
Senior	14 %	82 %	4 %
Non-freshmen	11 %	86 %	4 %
Entire cohort	13 %	84 %	3 %

Table 3: Percentage of words with positive, neutral, or negative sentiment per VADAR sentiment analysis of cohort by class

Table 3 shows that most classifications expressed similar sentiment in their responses with freshmen being more positive than freshmen. Breaking the cohort into their classes, Table 3 shows that number of positive words used in response is fairly constant, with sophomores being the least positive. The VADER analysis allows for neutral sentiment, so the negative sentiments does not necessarily follow the opposite of the positive sentiment. For example, students in all classes are equally negative in the sentiment expressed in their responses. It should be pointed out here that the seniors in this study are not well-represented, being only about 5% of the study population. Furthermore, recall that these seniors are not approaching graduation. They are seniors by their credit hours earned in university records. In fact, these seniors are enrolled in the very first course of a four-year engineering program of study. Students in the course have many more ECE courses to complete before graduation. Again, seniors in Table 3 are most likely students who were pursuing a different major for many semesters, or rarely, the student returning to university for a second undergraduate degree. Yet, this group of students express a relatively positive sentiment. One would like to think that these students, who possess more university experience than their colleagues, are comfortable in their assumed new academic direction. Freshmen also were the most positive in their responses. Could this be the exuberance of the beginning of a new stage in life which is their university matriculation?

Table 4 shows the compound score from VADER sentiment analysis and the valence averages from the AFINN, TextBlob, and ANEW2017 lexicons. The VADER compound score is a metric that calculates the sum of all the lexicon ratings which have been normalized between -1 (most extreme negative) and +1 (most extreme positive). The VADER compound score is often used to determine the overall sentiment of a statement, as most statements are composed of positive, negative, and neutral words. In general, a statement is regarded as positive when the VADER compound score is greater than +0.5. The statement is regarded as negative if the VADER compound score is less than -0.5. Otherwise, the statement is considered neutral. The scores reported for VADER compound score are calculated using the published VADER techniques and are optimized for social media. If a text selection has an overall positive or negative sentiment, the VADER compound score tends to saturate at +1 or -1 after a few dozens of words. The focus on tweets and short social media posts by VADER is apparent. Student responses were mostly much longer than a “tweet” and saturate near +1 (purely positive).

AFINN sentiment lexicon scores words as integer levels of negative and positive sentiment with -5 being the most negative and +5 being the most positive. The assumption is that zero would be “neutral” in the AFINN analysis. TextBlob also aggregates the sentiment of text to number in the range of [-1, +1]. ANEW2017 rates valence from +1 to +9. To aid in comparing the results of the different lexicons, results from AFINN and ANEW2017 were normalized to the range [-1, +1].

	VADER compound score	***Normalized AFINN valence***	***TextBlob valence***	***Normalized ANEW 2017 valence***
Freshmen	0.9995	0.018	0.139	0.077
Sophomore	0.9899	0.006	0.194	0.065
Junior	0.9987	0.018	0.091	0.084
Senior	0.9644	0.020	0.263	0.069
Non-freshmen	0.9996	0.013	0.161	0.074
Entire cohort	0.9999	0.015	0.153	0.075

Table 4: Sentiment analysis of cohort by class with scales -1 (purely negative) to +1 (purely positive) for VADER, normalized AFINN, TextBlob valence, and normalized ANEW 2017 valence.

Examining Table 4 shows that results mostly support the percentage measures in Table 3. The VADER compound score indicates that all populations would be considered overall positive in their response with the VADER compound score tending toward +1 when analyzing larger subject groups. AFINN and ANEW2017 results agree with the VADER percentages in Table 3 with sophomores being the most negative. TextBlob rates juniors as being the most negative class, and seniors being much more positive than the other classes. We note once again that the number of seniors in the study is quite small. The three analyses disagree in small ways between the classes, but all agree that class groups are positive, but only slightly so. The valence numbers are near zero which indicates a neutral emotion. Student responses are composed of many nouns and verbs that are technology-based. Such words are often not represented in the sentiment analysis lexicons, and if they are those words tend to be neutrally rated.

Table 5 show the percentages of words of each sentiment in the responses by different student majors as identified by the VADER sentiment analysis algorithm. Electrical engineering (EE) majors and were the most positive, and CmpE being the least positive. However, CmpE majors were the least negative. Students majoring in engineering were more positive than non-engineering and undeclared students. Non-ECE majors and Engineering-Undeclared students are noticeably more negative in their response than the other groups.

	Positive	***Neutral***	***Negative***
Electrical Engineering (EE)	14 %	83 %	3 %
Computer Engineering (CmpE)	8 %	90%	2 %
Electrical & Computer Engineering (ECE)	13 %	84 %	3 %
Non-ECE majors	10 %	85 %	5 %
Engineering majors	14 %	83 %	3 %
Non-engineering majors	9 %	89 %	3 %
Undeclared	9 %	88 %	3 %
Engineering-Undeclared	32 %	60 %	8 %
Entire cohort	13 %	84 %	3 %

Table 5: Percentage of words with positive, neutral, or negative sentiment per VADAR sentiment analysis of cohort by declared major

Table 6 shows the VADER compound score, and sentiment averages using the AFINN, TextBlob, and ANEW2017 lexicons for the study population by major. As before, the AFINN and ANEW2017 scores have been normalized to the range [-1, +1] for comparison. In this analysis, all populations would be considered very positive by VADER compound score classifications and the saturating nature of the VADER compound score is clear. TextBlob rates the non-engineering majors and undeclared students as the most positive, while AFINN and ANEW2017 rank the same groups as the most negative. Clearly, the lexicons used by the methods influence the results. All three lexicons that CmpE majors are the least positive of the engineering students, but not by a large amount.

The results reported here at not overly discriminating along any factoring of subjects. Students in all majors and classes responded with sentiment that is slightly more positive than neutral. The sentiments expressed by both university classification and declared majors are not very different in the big picture. Freshmen, sophomores, and juniors compose some 93% of the study group. These three classifications are all well-represented and the study group size is reasonable. Although the methods disagree a bit in their ratings, four of five methods indicate that freshmen appear to be bit more positive than the other classes. The reason for this is pure conjecture at this point. The authors attribute this result to freshman positivity to the “newness” of their situation. The freshmen are the newest university students. The freshmen have been on campus for a few days in the fall semester offerings and just starting their second semester in the spring semester offering. The excitement of the first week of university life among freshmen is palpable to anyone who has witnessed it first-hand.

	VADER compound score	***Normalized AFINN valence***	***TextBlob valence***	***Normalized ANEW 2017 valence***
Electrical Engineering (EE)	0.9997	0.016	0.168	0.079
Computer Engineering (CmpE)	0.9954	0.012	0.129	0.072
Electrical & Computer Engineering (ECE)	0.9999	0.015	0.150	0.076
Non-ECE majors	0.9675	0.016	0.170	0.066
Engineering majors	0.9999	0.015	0.147	0.077
Non-engineering majors	0.9382	0.011	0.223	0.050
Undeclared	0.9382	0.012	0.233	0.052
Engineering-Undeclared	0.8910	0.048	0.193	0.122
Entire cohort	0.9999	0.015	0.153	0.075

Table 6: Sentiment analysis of cohort by program with scales -1 (purely negative) to +1 (purely positive) for VADER, normalized AFINN, TextBlob valence, and normalized ANEW 2017 valence.

Differences in sentiment by the student's declared major is slight. Two of the methods (AFINN and ANEW2017) report that engineering majors, including EE, CmpE, EE+CmpE, and others (ME, ChE, AE, etc.), are more positive than non-engineering majors. While TextBlob goes the other way. All of the methods report in various ways and degrees that CmpE is less positive (but not necessarily more negative) than EE and others. One idea is that computer-related words in the lexicons, of which the computer engineers use more, carry a less positive score than words used by EE and others.

CONCLUSIONS

The authors teach an introductory discipline-specific course in electrical and computer engineering, in which most students have declared their intended major in either electrical or computer engineering. A small number of students enrolled have not declared their desired engineering program, or are currently declared as some other non-engineering major. Furthermore, while the course used in this study is a freshmen/first-year introductory course, the course is required of all electrical and computer engineering programs, and enrolls students of all classifications. The authors collected data on student career aspirations from almost 600 students over a four-year period with a question that demanded an open-ended, free-form prose response. Students answered in their own words. The student responses have been analyzed with textual data mining techniques and several sentiment analysis algorithms to ascertain most common thoughts and ideas and basic sentiment of the student responses. These results were analyzed with populations of the study group based on their declared major and university classification.

Electrical and computer engineering majors expressed aspirational job duties in terms of specific corporate entities, and mention broad concepts like *energy*, *power*, and *research* more often than non-ECE majors. Non-ECE majors mention non-electrical applications more than the EE+CmpE majors. Furthermore, the non-ECE majors were more likely to write about *job* and being an *engineer* as opposed to specific aspects or technologies of electrical and computing engineering, in agreement with other studies. Differences in popular ideas among the university classifications were more difficult to ascertain. In a very subtle way, freshmen expressed aspiration goals that were broad and more vague than "older" students who tended to mention more specific job duties or technologies.

When examining sentiment of the students' aspirations, freshmen and juniors appeared to have more positive career aspirations than sophomore. The ECE majors were more positive than any population who were not ECE majors. A future work is to extend the study of this population to examine the longitudinal results. The students reported these career aspirations very early – often the first few days on the university campus. Do these earliest indication of their career aspirations correlate with their ultimate longevity in the electrical or computer engineering program? The oldest data in this study belongs to students just now starting to graduate. A few more semesters need to elapse before persistence results can be evaluated.

REFERENCES

[Alp2008] E. Alpay, A.L. Ahearn, R.H. Graham, and A.M.J. Bull, "Student enthusiasm for engineering: charting changes in student aspirations and motivation," *Euro. J. Engr. Educ*., vol. 33, n. 5-6, pp. 573-585, 2008.

[Ben2015] D. Bennett, R. Kapoor, K. Rajinder, and N. Maynard, "First year engineering students: Perceptions of engineers and engineering work amongst domestic and international students," *Intl. J. First-Year in Higher Educ*., vol. 6, no. 1, pp.89-105, March 2015

[Bie2016] A.R. Bielefeldt and N.E. Canney, "Humanitarian Aspirations of Engineering Students: Differences between Disciplines and Institutions," *J. Humanitarian Engineering*, vol. 4, no. 1, pp. 8-17, 2016

[Bra2017] M.M. Bradley, & P.J. Lang, *Affective Norms for English Words (ANEW): Instruction manual and affective ratings*. Technical Report C-3. Gainesville, FL: UF Center for the Study of Emotion and Attention, 2017.

[Cav2007] M. Cavalli, L. Stanlake, and S. Tolbert, "Investigation of Retention and Perceptions among Freshman Engineering Students", *Proc. 2007 ASEE Midwest Sectional Conference*, 2007.

[Chr2014] R. Christensen, G. Knezek, and T. Tyler-Wood, “Student perceptions of Science, Technology, Engineering, and Mathematics (STEM) content and careers”, *Computers in Human Behavior*, Elsevier Press, 2014.

[Fel1995] R.M. Felder, G.N. Felder, M. Mauney, C.E. Hamrin, and E.J. Dietz, “A Longitudinal Study of Engineering Student Performance and Retention. III. Gender Differences in Student Performance and Attitudes, J. Engr. Educ., vol. 84, no. 2, pp. 151-163, 1995.

[Hut2014] C.J. Hutto and Eric Gilbert, “VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text”, Proc. Int’l Conference on Web and Social Media, 2014.

[Ita2016] M. Itani and I. Srour, “Engineering Students' Perceptions of Soft Skills, Industry Expectations, and Career Aspirations,” J. Prof. Issues Eng. Educ. Pract., vol. 142, no. 1, 2016.

[Lor2019] S. Loria, “TextBlob: Simplified Text Processing” available at *https://textblob.readthedocs.io/en/dev.* Most recently accessed April 27, 2019.

[Liu2010] B.Liu. “Sentiment Analysis: A Multifaceted Problem.” IEEE Intelligent Systems, vol. 25, no. 3, pp. 76-80, 2010

[NAB2004] National Academies Press, “The Engineer of 2020: Visions of Engineering in the New Century”. 2004.

[Nie2011] Finn A. Nielsen, “A new ANEW: Evaluation of a word list for sentiment analysis in microblogs”, Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts', pp. 93-98, May 2011.

[Pal2010] P.J. Palazolo, S. Ivey, and C.V. Camp, “Freshman Engineering Student Perceptions of Engineering Disciplines”, 2010 ASEE Southeast Section Conference.

[Pot2013] G. Potvin, Z. Hazari, L. Klotz, A. Godwin, R. Lock, J.D. Cribbs, N. Barclay, “Disciplinary Differences in Engineering Students' Aspirations and Self-Perceptions,” *Proc. 2013 ASEE Ann. Conf. and Expo.* Paper #7350, 2013.

[Ryn1988] S.L. Rynes, P.S. Tolbert, and P.G. Strausser, “Aspirations to Manage: A Comparison of Engineering Students and Working Engineers”, Cornell Univ. ILR School, 1988.

[Shi2003] V.A. Shivy and T.N. Sullivan, “Engineering students' perceptions of engineering specialties,” J. Vocational Behavior, vol. 67, pp. 97-101, 2005.