Gender Differences in Engineering: A Data-Driven Study

by

Shivangi Chopra

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Management Sciences

Waterloo, Ontario, Canada, 2022

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Karsten E. Zegwaard Associate Professor, School of Science, University of Waikato
Supervisor:	Lukasz Golab Professor, Dept. of Management Sciences, University of Waterloo
Internal Member:	Mark D. Smucker Professor, Dept. of Management Sciences, University of Waterloo
Internal Member:	Houra Mahmoudzadeh Assistant Professor, Dept. of Management Sciences, University of Waterloo
Internal-External Member:	Wojciech Golab Associate Professor, Dept. of Electrical and Computer Engineering, University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The gender gap in Science, Technology, Engineering, and Mathematics (STEM) is well known. Not only do fewer women apply to and earn engineering degrees, but also more women leave engineering programs and careers. Past studies have identified various reasons that affect female students' decisions to join engineering programs. Some of them include STEM interest, access to role models, and high school context. Further, labour market studies, which focus on later career stages, have found workplace experiences of engineering graduates to differ based on gender and drive female attrition. While the majority of studies on STEM recruitment are qualitative in nature or are based on small datasets collected using surveys and interviews, this thesis takes a data-driven approach towards studying gender differences in engineering. Moreover, since early career experiences can greatly affect subsequent career choices, this thesis investigates gender differences in early engineering careers, specifically in the co-operative education (co-op) form of work-integrated learning.

Our analysis is enabled by unique datasets from a large North American university with renowned engineering programs and mandatory co-op. We use standard statistical and text analysis tools to measure gender differences in (a) motivations, interests, and backgrounds of 33,763 applicants, and (b) co-op work experiences of 8,956 students in terms of their choices, opportunities, evaluations, and satisfaction. The goal of this thesis is to quantify the gender gap in engineering and provide data-driven insights into closing it.

While analyzing students' motivations behind joining co-op engineering programs, we find that female applicants are more likely to mention personal influences, a desire to contribute to society, and access to real-world work experiences. In addition, the unique characteristics of high schools that produce more female engineering applicants include: a) on average, female students from these schools outperform male students on standardized math tests, and b) applicants from these schools report more personal influence and a wider variety of interests, encompassing technology, arts, community, and travel. However, these applicants participate in fewer collaborative and competitive STEM activities.

Our analysis of students' co-op experiences shows that female students tend to apply to and fill slightly different positions than male students. While male and female students appear equally likely to obtain interviews and secure placements, female students seem to take more risks when ranking potential job opportunities and receive slightly higher performance appraisals. Nevertheless, male students appear to be perceived as more agentic and are more satisfied than female students from the very beginning of their careers.

The data-driven findings presented in this thesis may encourage female students to apply to engineering, as well as provide actionable insights to academic institutions and employers wishing to diversify their talent pool.

Acknowledgements

I would like to thank my supervisor, Prof. Lukasz Golab, for his constant guidance, especially with scientific writing, and the other members of my examination committee for their time and interest in this dissertation. I would also like to thank the members of the Work-Learn Institute and the Data Science Lab at the University, including Judene Pretti, Alex Jennings, Dave Drewery, Robert Craig, Dianne Bader, Melicaalsadat Mirsafian, Hannah Gautreau, and Abeer Khan, for helping me brainstorm certain ideas presented in Chapters 5 and 6.

Finally, I am eternally grateful to my husband for his unabated support, my mother for her selfless availability, my sister and father for believing in me, and my friends for encouraging me through the process.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Problem

The gender gap in Science, Technology, Engineering and Mathematics (STEM) education and workforce is well documented. Reports indicate that undergraduate engineering programs receive applications from and award degrees to only 22% women [326]. Statistics on earned bachelor's degrees in 2012 show that women were awarded 59% of degrees in the biological/biomedical sciences, but in math-intensive fields such as computer science and engineering, women were awarded only 18% and 19% of the degrees respectively [326]. Moreover, studies found women to be less likely to pursue undergraduate engineering degrees even when they are equally or more qualified than men. For example, only 23% of high school girls with high mathematics scores pursue engineering (ENG) degrees compared to 45% of boys with the same scores [144]. In fact, girls with high mathematical ability are less likely to enter engineering programs than even boys a lower mathematical ability (23% versus 39%) [144]. The under-representation of women continued at the graduate level, with women receiving only 19% and 23% of doctorates in computer sciences and engineering respectively [326]. Sources from various countries, including those in North America, report similarly low proportions of women enrollment in STEM and ENG since 1995 [144, 259, 325, 324, 244, 121, 236, 235, 122].

Not only do fewer women enroll in STEM degrees, but also a higher proportion of women than men leave STEM degrees and careers [144]. Studies analyzing students enrolled in engineering majors found women to be twice as likely as men to drop out of their programs, especially during junior years [113, 158]. Even among women who persist, satisfaction with the engineering major does not translate directly to pursuing a career in engineering

[4]. Even though women engineering students outperformed men academically [6, 53] and displayed a stronger commitment to pursuing their degree [141], 40% of women engineering graduates switched out of the field after graduation [157, 319]. In 2016, the engineering workforce of Canada comprised 83% men and only 17% women [259]. In addition, despite receiving 40% of the doctoral degrees in geoscience, women held less than 10% of full professorial positions [96]. These statistics indicate the existence of a gender gap in STEM.

1.2 Motivation

The first reason to study the gender gap is the social need for equality and inclusion. Women make up over 50% of the population and approximately 44% of the workforce, so it would stand to reason that women should constitute more than 20% of the engineers [231]. Moreover, the statistics presented above suggest that fewer female students apply to engineering programs despite qualifications and interest in STEM subjects, and more leave engineering programs and careers. Statistics like these have promoted the pursuit of gender equality and inclusion in STEM. Gender equality has been identified as one of the seventeen Sustainable Development Goals[1] that the United Nations (UN) aims to achieve globally by 2030. Particularly, global institutions including the UN, UN Educational, Scientific and Cultural Organization (UNESCO), UN Women, and UN Global Compact, have identified under-representation of women in STEM as a serious problem and have started various initiatives including the L'Oreal-UNESCO For Women in Science Programme, Organization for Women in Science for the Developing World, and STEM and Gender Advancement project, to make STEM accessible for women[2]. Furthermore, gender equality has gained political traction and support, with world leaders, including Madame Chancellor Angela Merkel and Justin Trudeau, addressing it as "just logical" and taking actions towards closing the gap[3]. Secondly, multiple reports state that increasing the participation of women in STEM education and workforce can counter the ongoing shortage of STEM workers[4]. Current studies have observed that the STEM workforce is not keeping pace with the needs of the labour market, in terms of size and quality [355, 35]. According

[1] https://news.un.org/en/story/2019/02/1032401, https://news.un.org/en/story/2019/02/1032221

[2] https://news.un.org/en/story/2019/02/1032401

[3] https://www.zeit.de/politik/deutschland/2019-01/angela-merkel-chancellor-cdu-feminism-interview, https://www.theguardian.com/world/2015/nov/04/canada-cabinet-gender-diversity-justin-trudeau

[4] https://www.usnews.com/news/best-countries/articles/2018-08-23/americans-think-they-have-a-shortage-of-stem-workers

to one report, the United States needs to increase the number of STEM graduates by 34% (i.e., produce approximately one million more graduates) to match the demand forecast for STEM professionals and maintain the lead in the global economy [355, 290, 152]. Developed nations, including the United States and the United Kingdom, fear that the shortage of STEM workers will lead to significant economic implications and loss of competitiveness [35, 292, 355]. Additionally, since the world is heavily dependent on technology-based goods and skills, the shortage of appropriately skilled workers is a threat to the global economy, health, and security [292]. Thus, participation of women in science has been deemed necessary to compensate for this deficit, make urgent technological advancements, and improve global economy. On February 11th, 2019, the UN Secretary-General Antonio Guterres announced that the participation of women in science was vital to achieving the 2030 Agenda for Sustainable Development[5]. He elaborated that with so many areas needing attention, the world cannot afford to miss out on the contributions of half its population.

Thirdly, participation of women in STEM will exclude bias from and increase diversity in product design. The World Economic Forum's Global Gender Gap Report shows that only 22% of artificial intelligence (AI) professionals globally are women. However, diversity of those working on AI solutions has been identified as a crucial element in ensuring that they are free from bias[6]. Moreover, diversity helps in adding different perspectives through which problem definition and problem-solving can occur [29, 151, 152]. For example, when the first voice recognition systems were designed, they were calibrated to men's voices and the voices of women were not recognized [213]. Additionally, the lack of women in the engineering workforce not only impacts the design of products and services that are used by women, but also impacts the safety and efficacy of such products used for and by women [152, 290]. For example, when the first automobile airbag systems were designed, they were designed around the specifications of a man's body and the lives of many women were lost [213]. Such limitations in the design of products and services might be reduced by having more women involved in the engineering process.

Moreover, since women comprise over 50% of the population, having a more diverse engineering workforce will provide a better match of products and services to this diverse customer base [151]. Past research points out that the lack of involvement of women in the design of technology will continue to result in the production of technologies that do not respond adequately to women's concerns [152, 119, 81]. For example, keeping in mind that women purchase 65% of all cars and influence the purchase of approximately 80% of all car purchases, the Volvo automobile company unveiled its first concept car designed by

[5] https://news.un.org/en/story/2019/02/1032221

[6] https://www.weforum.org/reports/the-global-gender-gap-report-2018

a team of all women engineers. The car includes features that might be more attractive to women such as no gas cap, compartments for handbags, and a swing-out seat for ease of entry [307]. Arguing that gender diversity helps businesses to perform better, UN Women and UN Global Compact are working together to encourage the private sector to sign up for the Women's Empowerment Principles[7].

Lastly, fewer women in engineering leads to fewer women faculty, entrepreneurs, and professionals, especially at visible positions. The lack of women role models creates a lack of perceived similarity and belonging in young girls and inhibits them from pursuing engineering [273, 81, 55], making the problem circular. Due to the above reasons, there is political, social, economic, and industry pressure to increase the number of women in STEM [292, 35].

1.3 Existing Work

Numerous studies have analyzed various reasons behind the gender gap in STEM. A major focus has been to identify reasons behind the "leaks" in the pipeline [52, 69, 85, 165, 153]. To that end, much work has been done to investigate (a) why fewer women join engineering programs [290, 270, 85, 339, 314, 30, 82, 299, 340, 227, 342, 337, 102, 278], and (b) why more women than men leave engineering programs and careers [113, 209, 137, 302, 69, 272, 90, 96, 285, 293, 269, 182, 303, 190, 157, 132, 174, 4, 48, 26].

Gender differences in engineering major choice and persistence have been studied from various perspectives [16, 85, 339, 136, 191]. These include differences in students' academic performance in STEM courses [342, 333, 309], interest in the field [308, 169, 208, 278, 102], confidence in STEM ability [113, 224, 250, 283, 351], parent's socio-economic backgrounds [113, 31, 327, 134, 228, 192], sources of influence (including parents and teachers) [57, 287, 353, 220, 206, 143, 315], need for social belonging [314, 302, 137, 285], career perceptions [339, 100, 87, 102, 13, 221, 125], location and resources of their high school [180, 313, 220, 249, 245, 82], and belief in the pro-male STEM stereotype [227, 169, 225, 339, 309, 200, 104]. While these studies have generated insightful findings, they are limited by relatively small sample sizes, which typically originate from surveys, interviews, and longitudinal studies. The consensus report prepared by the National Academy of Engineering in 2018 recognized this limitation [243]. They recommend that researchers "work with institutions of higher education ... and build on administrative data resources to establish a better empirical foundation for research on the educational and career paths of engineers ... for

[7]https://www.empowerwomen.org/en/weps/about

the purposes of building a robust evidence base to inform policy" [243]. Other studies have made similar observations [339, 69, 85].

Moving on, students' career paths post-graduation have been studied extensively using both qualitative and quantitative methods. In order to understand why more women leave engineering careers, researchers have studied gender differences in opportunities received (in terms of hiring, salary, and promotion) [157, 49, 303, 23, 257, 103, 354], job attribute preferences [317, 118, 339, 48, 132, 138], satisfaction [155, 26, 184], perceived competency [293, 269, 96, 182, 232], and working conditions [157, 285]. Even though most of these studies are based on later careers, researchers suggest that early career experiences drive attrition more than other factors [132, 157, 174].

1.4 Our Work

Our research contributes to the efforts mentioned above with novel data-driven analyses of STEM's educational pipeline. The goal of this research is to measure the gender gap in engineering by quantifying gender differences in engineering applicants and students and suggest data-driven insights towards closing this gap. We do this by applying statistical and text analysis tools to large and unique datasets from the engineering faculty of a large North American university.

As seen in Figure 1.1, first, we analyze the secondary education stage and answer research questions related to the interests and backgrounds of engineering applicants. In contrast to prior qualitative work, our research is enabled by access to a unique dataset of over 30,000 undergraduate engineering applications to the engineering faculty of a large North American university. To the best of our knowledge, there has been no prior large-scale data-driven analysis of engineering applicants' motivations, interests, and backgrounds. *Most of our work on this stage has been published [60, 67, 64], and some of it is under review.*

Next, we analyze gender differences in engineering students' co-operative work experiences (Figure 1.1). Co-operative education, or co-op, is one of the nine types of work-integrated learning (WIL) [247]. Co-operative Education and Work-Integrated Learning Canada (CEWIL) defines WIL as "a model and process of curricular experiential education which formally and intentionally integrates students' academic studies within a workplace or practice setting" [247]. Students enrolled in co-op programs alternate between periods of academic study (at the university) and relevant paid work experience (at their employers' workplace) [247, 74, 140]. Since co-op jobs represent the first STEM work experiences for many undergraduate STEM students, they have the potential to affect future career choices, including attrition. Therefore, in contrast to prior work that analyzes later careers, our research uses co-op work term experiences to investigate gender differences in early career experiences of engineering students.

Our analysis is enabled by unique data extracts containing job search, hiring, and evaluation data for nearly 9,000 students enrolled in the undergraduate engineering co-op programs of a large North American university. We use standard data analysis techniques to quantify gender differences in these co-op work experiences, specifically in terms of the opportunities received, choice, perceived competency, and satisfaction. To the best of our knowledge, this is the first work to study co-op education from a gender perspective. In addition, while past studies on gender differences in job search are based on surveying particular employers and candidates and thus are prone to incomplete information or bias, access to recruitment information for all students competing for all available jobs in a closed (co-op) labour market gives us a unique opportunity to study gender differences in hiring as well as the internal decisions along the process. *While most of our work on this stage has already been published [63, 66, 65, 67], some of it is forthcoming [62].*

To summarize, in contrast to past studies that use small samples collected through surveys and interviews, we conduct a holistic, large-scale, secondary data analysis of unique datasets to understand gender differences in engineering students' interests and experiences before and after joining co-operative undergraduate programs.

1.5 Process Overview

Figure 1.2 illustrates STEM's educational pipeline, along with the components of every stage that we analyze in this study. Our analysis of the secondary education stage is enabled by access to datasets corresponding to the two components shown in Figure 1.2: application forms and aggregate data about high schools. The first contains data extracts with over 33,000 undergraduate applications – both accepted and rejected – to the engineering faculty of a large North American university with mandatory co-op programs. In their applications, prospective students are required to describe why they are interested in studying engineering, specifically at the university, and provide other relevant information such as their reading interests, extracurricular activities, and the last academic institution they attended. The second component contains aggregate demographics and academic performance statistics for all public high schools in the province of Ontario in Canada, where this university is located.

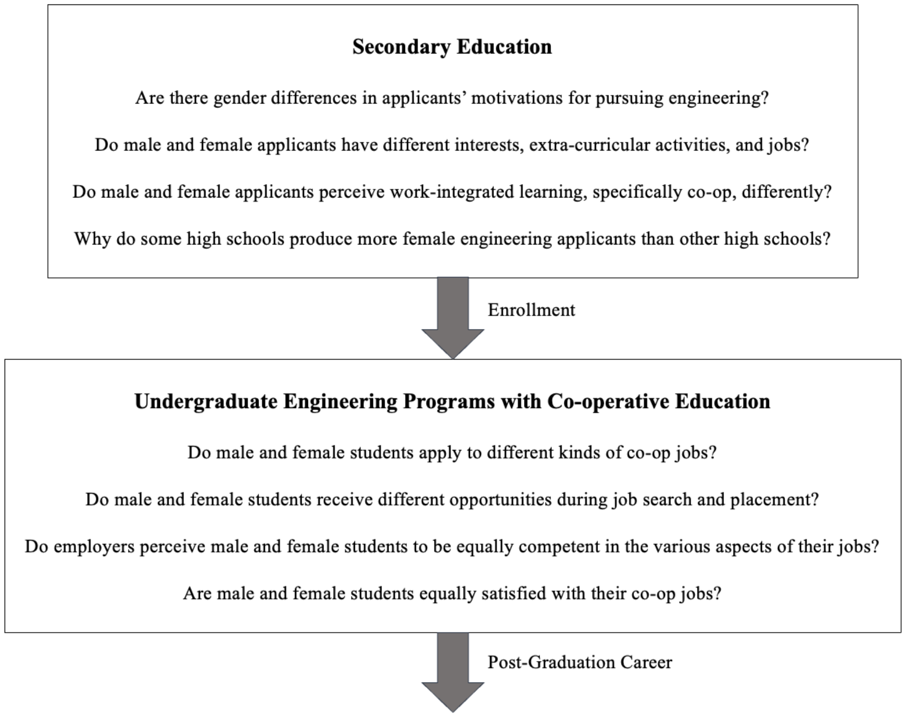

Figure 1.1: Research questions

Combining these datasets, our goal is to determine whether male and female applicants to engineering programs have different interests and backgrounds. Particularly, in Section 5.1 of Chapter 5 we examine gender differences in engineering applicants' motivations and interests *(published [60, 67])*, in Section 5.2 we examine gender differences in applicants' perceptions of co-op *(published [64])*, and in Section 5.3 we identify unique characteristics of high schools that produce many female engineering applicants. The specific research questions are mentioned in Figure 1.1.

Moving on, accepted applicants join the undergraduate engineering co-op programs at the university. As mentioned above, co-op is a form of WIL that combines experiential education with academic studies [247, 74, 140]. It includes both academic study terms and paid work experience, referred to as co-op placements, work terms, or internships. It provides new learning opportunities for students, a talent pipeline for employers, and a recruiting tool for institutions [74, 116, 98, 318, 140]. Engineering co-op programs were introduced in North America in 1906 [140]. Since then, co-op education has been adopted by many disciplines worldwide, with at least 116 institutions offering co-op programs in

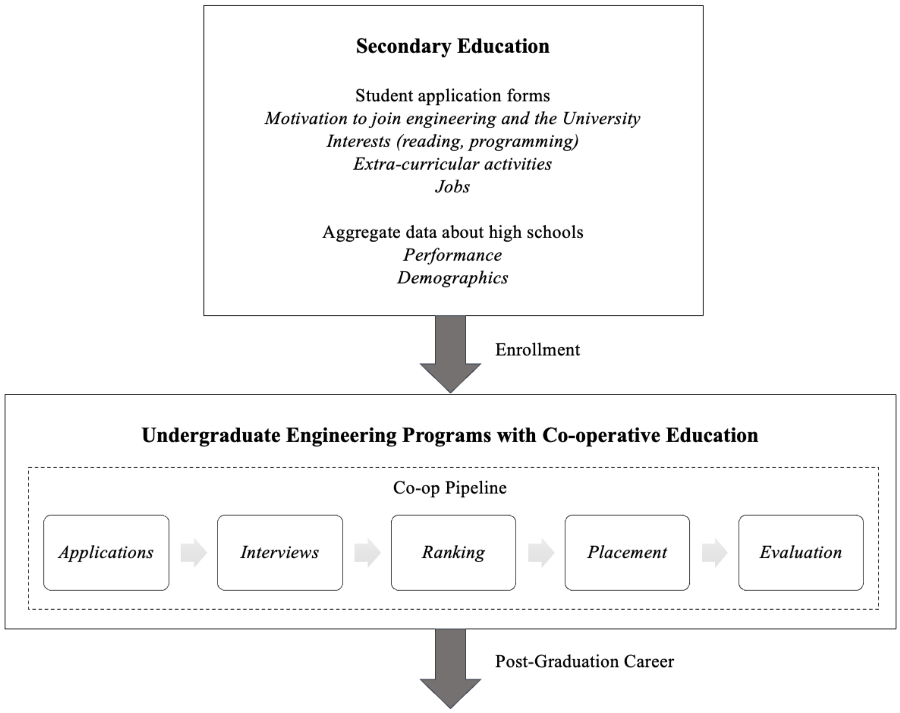

Figure 1.2: Components of the educational pipeline

Canada alone [247]. For many students, co-op work terms are their first career experiences in the workplace.

Access to unique data extracts containing records of students' and employers' activities in the co-op system enables our analysis of early careers. These data extracts were not collected through surveys or interviews. Instead, they contain detailed information about each stage of the co-op pipeline (Figure 1.2), which proceeds as follows. Initially, employers participating in the co-op process submit job descriptions to the university, and any student can apply to any job. Next, employers interview selected candidates and rank the ones they are willing to hire, by either offering them the position or shortlisting them for it. After the students have responded to these ranks, the university follows a matching process to assign students to jobs, with the goal of minimizing the sum of the student and employer ranks. Ideally, as many students and employers as possible should get their top choice, but some may hire or be placed at their second or third choice depending on the level of competition, and some students or employers may not be matched at all. Finally, at the end of a four-month work term, students and employers evaluate each other. Engineering programs at the university mandate all enrolled students to participate in the above mentioned co-op process and complete at least five work terms before graduation.

By examining the data generated from all stages of the co-op process, our goal is to determine whether students experience any difference in opportunity, evaluation, satisfaction, and choice due to their gender (Figure 1.1). To that end, each section of Chapter 6 analyzes a particular stage of the co-op pipeline. Particularly, Section 6.1 examines gender differences in the application stage of the co-op pipeline, where we determine whether male and female students differ in the number and kind of co-op jobs they apply to *(published [66])*. Section 6.2 analyzes the interview and ranking stages, where we inspect gender differences in the interviews, ranks, and offers received and students' responses to them *(published [66, 62])*. Section 6.3 analyzes differences in the jobs held by male and female students *(published [66, 67])*, Section 6.4 inspects gender differences in the numeric as well as textual performance evaluations *(published [63, 66, 65])*, and Section 6.5 analyzes gender differences in students' satisfaction with their co-op work terms *(published [66])*.

To recap, in this study, we present a data-driven analysis to gain a holistic understanding of gender differences in engineering. In particular, we analyze interests and backgrounds of engineering applicants and early career experiences of undergraduate students to understand how to attract and retain female students to engineering programs and careers. Using real datasets from a large North American undergraduate institution, our goal is to measure the gender gap and suggest how to close it. Analyzing unique datasets allows us to verify some results of previous studies as well as obtain new insights.

In the remainder of this report, Chapter 2 summarizes the related work; Chapter 3 discusses the datasets that we have and the limitations of the research (Section 3.4); Chapter 4 provides details of the methods used to analyze the data; Chapters 5 and 6 summarize the results and insights obtained from the research on gender differences in students' experiences during secondary education and co-op work terms, respectively; and Chapter 7 highlights the conclusions of this research.

Note: permission for this secondary data analysis was granted by the university's office of research ethics (application number 40471).

Chapter 2

Related Work

This section lists prior work on gender issues in STEM education and careers. Particularly, we summarize past research on (a) why fewer women join engineering programs (Section 2.1) and, (b) why more women than men leave engineering programs and careers (Section 2.2). Since researchers have identified these as the two largest "leaks" in STEM's educational pipeline leading to the under-representation of women in STEM education and careers [165, 153], this section summarizes past work related to these topics.

2.1 Why do fewer women join engineering programs

Various sources from different timelines and countries in North America report low proportions of female enrollment in engineering programs [311, 326, 144, 259, 325, 324, 244, 121, 236, 235, 122]. Even after spending over 2.8 billion dollars on numerous outreach programs and recruiting initiatives [7], this proportion has been at a constant low [326]. To understand why fewer women than men apply to engineering programs, we start by summarizing past research on their background, especially differences in their socio-economic and high school background. Then, we summarize past research on gender differences in ability, performance, confidence, and interest in STEM subjects. As discussed below, these differences may have stemmed from various overlapping factors, including the widespread stereotypes regarding the pro-male STEM ability bias, difference in support and encouragement from parents, peers, and teachers, access to relatable role models, lack of social belonging, low perceptions about the impact of STEM careers, and mismatch of life priorities with STEM careers. While most research on STEM recruitment is qualitative in nature and is based on small populations of either high school students or students already enrolled in engineering programs, some, especially those related to demographics and performance, are quantitative and apply statistical and data mining techniques on census and other summary data. Below, we summarize past work on under-representation of female students in engineering programs from various perspectives.

2.1.1 Demographics

Studies found that race, ethnicity, gender, and parents' economic status and education level affect students' decisions to apply to college [270, 327, 31, 134]. In addition, demographics also affect the choice of major, especially in STEM [16, 133, 313, 206]. A random forest model built on data of 13,000 high school students considered both student-level (demographics, family background, and interest and performance in STEM) and school-level factors (urbanicity, resources such as the number of STEM courses offered and teacher education level, and percentage of students eligible for free lunch) found gender to be the most important variable in predicting engineering major choice [313]. Moreover, a study that used a logistic regression model to differentiate between white males and others belonging to different races and gender found that white male students had a higher STEM self-concept and were more likely to be interested in STEM careers [270].

Women from families with a higher socio-economic status were more likely to pursue a STEM degree [191, 230]. A longitudinal study with 87 male and 34 female engineering students found female engineering students' parents to be more educated or trained in science or technology [113]. Moving on, past studies that examined student backgrounds, including gender, socio-economic status, ethnicity, and country, to understand their effects on participation in *co-op programs*, found fewer female students to enrol in co-op programs at the university level [95]. While these studies focus on student demographics and are based on census and other longitudinal surveys, our work focuses on identifying gender differences in students' motivations and backgrounds based on what they say in their engineering admission forms.

2.1.2 High School Context

Researchers note that less work has been done to understand the effect of a student's surroundings, especially their high school contexts, on their post-secondary choice of major [136, 191, 256, 105]. For example, Lee [191] suggested that more work is needed to understand situations "in which participants are acted upon by a surrounding system and have little agency to change their course". In response, recent works studied the relationship between a student's choice of engineering major and the high school they attended [180, 313, 220]. Researchers considered the high school's (a) *demographics* (specifically, the high school's socio-economic status, calculated using the percentage of students eligible for free or reduced-price lunch [248, 203, 249, 313], and, its district's socio-economic status, calculated using median income and proportion of adults with bachelor degrees [181, 180, 249]), (b) *resources and experiences* (school size, number of advanced STEM courses offered, number and quality of STEM teachers, counsellors, budget to organize extracurricular activities and college preparation programs, disciplinary climate, and performance averages [327, 83, 220, 313, 131, 12, 105, 313]), and (c) *location* (including urbanicity and distance from post-secondary institutions [237, 313, 82]).

It was found that students' choices to apply to college and their choice of major are affected by school experiences, which are closely tied to school resources and the demographics of the neighbourhood where the school was located [327, 12, 313]. Students from schools with a low socio-economic status were found to be at a disadvantage when it came to post-secondary engineering matriculation, achievement, and persistence [248]. In addition, studies found that rural students were less likely than suburban or urban students to choose an engineering major [237, 313]. Furthermore, an inverse relationship was found between the proportion of engineering graduates from a high school and its distance from the closest post-secondary engineering institution [82]. *All these relationships were found to be stronger for underrepresented groups such as gender and race.* Women scientists reported that their school experiences, especially in peer groups, were crucial to the development of their interest and curiosity in science [208, 188, 80, 301]. While most of these findings are based on descriptive statistics and regression analysis of census data, our work is the first to apply text analysis on engineering applicants' admission forms to infer differences in their backgrounds.

2.1.3 Ability and Performance

While early attention focused on directly measurable traits, for example, the size of the head as an indicator of brain size and indirectly a measure of intelligence, scientists eventually found evidence otherwise and discarded the theory that women's biologically-driven intellectual inferiority contributed to their low participation in STEM [36, 37, 299, 277, 161, 69]. In fact, studies found that girls were more likely than boys to possess both high mathematical and verbal abilities, whereas boys were more likely to demonstrate higher mathematical abilities relative to their verbal abilities [340, 339, 329]. Researchers speculate that these multiple cognitive strengths afford female students a wider variety of career choices, suggesting that the gender gap in STEM programs is due to factors other than a difference in technical aptitude and ability [339, 338, 340, 342, 337].

Since past studies have found intent to pursue STEM programs to be directly affected by high school math achievement [342, 338, 161], we now consider gender differences in math performance. Multiple studies have applied statistical and data mining techniques to understand the effect of gender on STEM performance of high school and first-year engineering students [276, 114, 159, 185]. While some of these studies found that girls earned higher grades than boys even when they were not interested in science and did not participate in science-related extracurricular activities [311, 332, 113, 46, 241], others found small but statistically significant differences in performance favouring boys, from as early as kindergarten [198, 271, 333, 334, 309, 160, 161, 69]. Not only did boys outperform girls on complex programming tasks [309] and take less time on their homework [114], but they also outnumbered girls by approximately 4:1 and 3:1 in the top 0.01% of the distribution for the math subtests of the SAT and ACT, respectively [333, 334].

With no difference in technical ability (due to biology), Wang & Degol [339] speculate that the socio-cultural factors, such as stereotypes about gender differences in STEM ability and how they manifest in students' educational experiences, may have lead to gender differences in self-ability beliefs, performance, and career ambitions in STEM. For example, surveys of 100-200 students reveal that girls may be avoiding STEM courses and careers because they not only erroneously believe that innate intelligence is needed for success in these fields, but also that they belong to a group that is less likely to possess those abilities [194, 225, 339]. The stereotype that men are better at math and science is so pervasive that children as young as six subscribe to it [227, 309, 216, 169]. High school textbooks mentioning the work and pictures of more men scientists in comparison to women might propagate a similar gender bias [25]. Moreover, researchers that observed interactions found that parents and teachers underestimate girls' math ability relative to boys, despite having similar grades [315, 24, 201, 321, 315]. Not only do they encourage boys more often in math and science pursuits [315], but they also tend to attribute boys' successes in math more to ability and failures in math more to lack of effort, while the opposite is believed to be true for girls [320]. Since female students are more likely than male students to internalize the feedback they receive [239], it may lower their self-efficacy beliefs and also their performance in STEM courses and activities [240, 222]. Additionally, the prevalence of gendered STEM stereotypes may trigger stereotype threat in women, making them perform negatively and confirm to these stereotypes [86].

As a result of the above, female students may feel less capable, perform poorly, and in turn, be less interested in math and science [338]. Consequently, they are less likely to participate in STEM extracurricular activities or take advanced STEM math and science courses [24, 69, 304, 33, 46, 21]. To this end, a study that tested 206 first-year college students on high school STEM concepts, particularly electrical circuits, noted that it was not gender, but factors such as prior knowledge, interest, and experience that dictated performance, and in turn, choice of a STEM career [50].

That being said, other studies have observed that students' performance may not be the only factor behind the under-representation of female students in engineering programs. Quantitative studies based on summary statistics have found female students to be less likely to choose a STEM program, regardless of their mathematical ability [144]. While only 23% of female students with high scores in math choose a STEM program, 39% of male students with low scores opt for them [144]. These studies may indicate the existence of reasons beyond ability and performance for the low proportion of female students in STEM programs. Thus, our work focuses on students' motivations for studying engineering, and their non-academic experiences and backgrounds.

2.1.4 Interest

Gender differences in interest in math and science is one of the leading causes of low female enrollment in STEM [342, 339]. Many studies have found that from a very young age, girls are less interested in math and engineering, thus reducing their participation and performance in related subjects [1, 169, 208, 209, 42, 278, 308, 309]. While men show a greater interest in practical and investigative subjects and are more interested in physical sciences, mathematics, technology, and engineering, women show a stronger artistic and social bent and are more interested in biology, social sciences, environment, health, and medicine [68, 278, 308, 1, 2, 169]. In addition, female students are more interested in in-class activities, whereas male students are more interested in extracurricular activities and competitions [128, 46, 169, 24].

While most female students shy away from calculus and physics [24], those who are interested are more likely to pursue engineering [278]. On average, students who enrol in engineering programs take more advanced math and physics courses in high school [327, 24]. Since women are less interested in math, and in turn, are less likely to elect advanced math courses, it reduces their chances to enroll and persist in math-intensive STEM programs [24, 69, 208, 209, 76, 21]. Sadler et al. [278] conducted a retrospective study of 6,000 college students and noticed that men's interest in a STEM career was stable over their high school years, but women's interest declined near graduation.

As discussed in Section 2.1.3, female students' low interest in STEM subjects and professions may stem from the prevalent pro-male STEM stereotype [200]. A survey of 437 students revealed that not only do girls find science difficult to understand [309], but they also find it more suitable for boys [169]. Nations with higher proportions of women in post-secondary science courses and careers are less likely to explicitly endorse the stereotype that science is a masculine profession [227]. Alon & DiPrete [2] found that the first choice of engineering applicants is affected by gender stereotypes about the field (in particular skill requirements, work conditions, and gender composition of fields), however, the rest are motivated by interest. Bystydzienski et al. [42] surveyed 24 interested and high-achieving female high school students to find that they decided against pursuing STEM programs due to lack of financial and social support. Furthermore, a survey of 132 high school students [143], four high schools [220], and 20 university students [206], found personal influences, including parents and teachers who served as role models and provided resources and encouragement, to increase students' interest in STEM. Additionally, female students were more likely to choose STEM programs if their parents and teachers had a positive perception of STEM careers [143, 139, 353, 230, 131, 228].

In addition, past research suggests that women's lack of interest in STEM subjects and careers may be related to their low math task values, that is, the degree to which they believe that these tasks are *worth pursuing* and the value they attach to them [340, 337]. Girls with high math achievement and little motivation in pursuing a STEM occupation are far less likely to obtain a science degree than individuals with average math skills and high interest in science [312]. Being capable in math and science does not necessarily mean that an individual will enjoy STEM-related activities or want to pursue a STEM career [339, 102, 100, 337, 13]. Researchers state that choosing a career requires both the ability to pursue a career as well as the motivation to employ that ability [97, 102, 100, 202, 337, 13, 342]. Since women possess cognitive strengths in both verbal and math ability, their abilities take a backseat to their interests and values [339, 341]. Thus, even among mathematically talented individuals, women are more likely than men to pursue and make accomplishments in non-STEM careers [252].

Low math task values may be a result of the widespread stereotypes regarding the applicability of math and math-intensive careers, especially to communal goals [88, 87, 99]. Since women exhibit altruistic tendencies from a very young age [308, 169], they may overlook engineering careers because of their apparent inconsistencies with communal goals of collaboration and helping others [88, 87, 99, 124]. Eccles [99] conducted a longitudinal study on 1500 participants from sixth grade to adulthood and found that the main source of gender differences in entry to STEM programs was not gender differences in mathematical ability, but differences in inclination towards society-oriented jobs. Additionally, they found that women who aspired towards STEM careers placed a lower value on society-oriented job characteristics than their female colleagues who did not aspire to STEM careers [99]. Diekman et al. [87] interviewed 360 students from STEM and non-STEM fields and found that the popular perception of STEM careers was that they impeded the pursuit of altruistic goals. In fact, even within engineering, women obtained degrees in the obviously altruistic fields of biomedical and environmental engineering at higher rates than in mechanical or electrical engineering [48].

Essentially, the lack of awareness regarding how STEM careers lead to societal improvement and, in turn, the perceived mismatch of STEM careers with women's career aspirations inhibit them from pursuing STEM degrees. Programs that increase students' STEM task values (for example, scientific narratives that provide memorable real-life applications of the subject [177]), alter women's perceptions of STEM, and raise awareness about its applicability, especially to communal goals, were shown to increase female students' interest in STEM careers [338, 13, 221, 125, 207, 88, 301].

As seen above, gender differences in STEM interests are reinforced through a continual process of decision-making, experiential outcomes, and expectations of others [339]. In addition to the above, women may not be interested in STEM as they do not see its careers in line with their lifestyle values of work-life balance and family [102]. Cheryan & Plaut [55] who surveyed 33 female and 30 male students found that female students were reluctant to pursue male-dominated fields such as computer science because of a lack of perceived similarity and belonging and the prevalent discrimination that comes along with it. Similarly, Matusovich et al. [221] who interviewed 6 female and 5 male engineering students found that female students had more difficulty than their male peers in connecting their personal identities to engineering. Since women prioritize fit, personal values, and lifestyle goals while making career choices and view STEM careers as incongruous with these goals, fewer women than men may pursue engineering careers [339]. While past studies survey or interview school and university students to understand the role that interest plays in fewer female students pursuing STEM programs, our study conducts a large-scale text analysis of the interests reported by engineering applicants on their admission forms.

2.1.5 Confidence

In addition to past math achievement, students' math self-efficacy beliefs play an important role in their decisions to choose STEM programs [19, 279, 342, 327, 313, 314, 338]. Correspondingly, studies note that women have a lower math ability self-concept than men [102, 188, 201, 271, 305, 338, 339, 250, 214]. Researchers surveying 391 students found that, despite equal academic performance, male students held a better attitude towards computer and science-related tasks and were more confident with them than their female peers [33]. This lack of confidence may lead women to either not choose STEM programs, or even if they do, mostly choose non-math STEM programs [102]. Nevertheless, female students who do join engineering programs start with lower self-confidence [224, 250], a lower self-concept about their performance in STEM courses [109, 283, 284], lower self-efficacy [20, 32, 40, 73, 205, 295, 146, 193, 351], and greater anxiety about their preparation [113, 39] in comparison to their male peers. As students progress through the curriculum, poor grades lead to greater drops in self-esteem for female students [78].

As discussed before, female students' STEM identity and math ability beliefs are greatly affected by the gender stereotypes held by their parents, teachers, and peers [320, 321]. In interviews with 17 female engineering students, Smith [290] observed that influences from family or friends, including the support and encouragement received, played a pivotal role in helping them build self-confidence in their math and science ability. Interventions that provided access to diverse role models, in-person (to 33 students) or through videos (to 41 students), not only observed statistically significant increases in female students' success beliefs in STEM, but also observed an increased sense of compatibility and raised awareness of STEM career possibilities [57, 353]. Youth with peer groups who encouraged, endorsed, or exemplified high math and science achievement had higher math and science motivation [188], took more math courses [80], and were more likely to see themselves as future scientists [301].

In addition to the ingrained pro-male ability bias, researchers suggest that female students' fixed mindset and fear of failure may reduce their competence beliefs in STEM subjects and withhold them from applying to engineering programs [42, 110, 337, 338]. Not only do female students consider math and science to be male domains [106, 107, 110], but they also believe that their math ability is fixed and cannot be improved [337, 110]. Studies suggest that women may be more likely to pursue math-intensive STEM fields if greater emphasis is placed on effort rather than intelligence [339, 328]. Additionally, since technology is both materially and symbolically masculine [107] and female students are outnumbered by male students in STEM activities and competitions [149], female students may lack a sense of belongingness and compatibility with STEM, further lowering their competence beliefs in STEM subjects and careers [56, 54, 162].

Summary: Studies utilizing census data suggests that fewer women than men from low socio-economic households and schools join engineering programs. Even though prior work ruled out gender differences in technical ability as a reason behind the gender gap in engineering programs, female students were found to have lower STEM performance, interest, task values, experience, and confidence in their background knowledge than male students. These limitations were caused by various factors, including lack of support and encouragement from parents, teachers, and peers, lack of social belonging, lack of accessible role models, doubts about the social impact of STEM careers, life priorities that did not align with STEM careers, and most importantly, the internalized and widespread pro-male STEM stereotype.

While the studies discussed in this section generated insightful findings, they are either based on quantitative analyses of summary statistics or qualitative studies limited by small sample sizes (typically originating from surveys, interviews, and longitudinal studies). To the best of our knowledge, our study is the first to conduct a large-scale quantitative text analysis of the motivations, interests, and backgrounds of male and female engineering applicants as well as provide data-driven insights into their differences. It is also the first to differentiate between the expectations of male and female applicants from *co-op* programs. Additionally, as can be seen above, the majority of past studies focus on female students' individual considerations of joining engineering programs. Therefore, our study contributes to the growing yet small body of work that studies the effect of female students' high school context on their decision to pursue an engineering major [136, 191, 256, 105].

2.2 Why do more women leave engineering

While fewer women join engineering programs, more women than men leave engineering programs and careers [326, 259]. Even though undergraduate women in engineering indicate the same intent to persist in their degrees as men [73, 141], they leave the program at a much higher rate [113]. Among those who join engineering careers, attrition rate is higher for women than for men [259, 96, 171]. While female professors hold less than 10% of full professorial positions [96], they leave the profession 2.5 years sooner than male professors [171]. In this section, we discuss reasons that contribute to women leaving engineering, after having joined the program. There is a wealth of both qualitative and quantitative literature on understanding female students' attrition from STEM programs and careers, including gender differences in interests, evaluations, hiring practices, and promotion opportunities. Based on these studies, we gather four overarching reasons behind higher female attrition from engineering programs and careers. These are gender differences in (a) perceived competency, (b) opportunity, (c) satisfaction, and (d) choice. Unless specified otherwise, studies related to careers analyze the later stages of men's and women's post-graduation careers.

2.2.1 Perceived Competency

One of the reasons why women leave STEM programs and careers is because they are perceived to be less competent than men, not only by their peers and professors, but also by their prospective and current employers [293, 182, 4, 285, 346, 267, 294, 77]. While very few studies found STEM men and women to receive similar evaluations [331], most others found them to receive differential treatment, evaluation, and feedback in all study and work environments [293, 182, 316, 267, 232, 190, 96].

A double-blind experiment with 127 faculty members found both male and female faculty to be biased towards male students [232]. This was confirmed by interviews with 11 female engineering students where researchers analytically coded the interview transcripts [90]. Additionally, a manual coding of 1,224 recommendation letters for graduate studies in geoscience revealed that female applicants were only half as likely to receive excellent versus good letters compared to male applicants. Male and female recommenders were equally likely to display this bias [96]. Knobloch-Westerwick [182] set up an experiment with 243 scholars who rated conference abstracts ostensibly authored by men or women. Descriptive statistics indicate that publications from men were associated with greater scientific quality, particularly if the topics had traditional masculine themes. In addition, collaboration interest was highest for men authors working on stereotypically masculine topics. Again, respondent gender did not influence these patterns [182]. Terrell et al. [316] found that on the open source software website Github, women's contributions were accepted more often than men's. However, for contributors whose gender was identifiable, men's acceptance rates were higher [316]. The descriptive statistics reported by the authors suggest that although women on GitHub may be more competent overall, bias against them may exist nonetheless.

A study where 194 technology professionals rated *hypothetical* interns on competence, intelligence, and potential field issues found that men were rated more highly than women [267]. Qualitative coding of their free-text recommendations suggested that women with ability issues were viewed as having lower field aptitude than men with ability issues. In addition, men and women with interpersonal issues were given similar aptitude ratings, but men were dissuaded from seeking help while women were expected to find mentors and control their emotions [267]. Similarly, science faculty rated female applicants for a laboratory manager position as less competent and hirable than male applicants, despite having the same application materials [232]. Lee and Huang [190] who studied venture capitalists' evaluations of entrepreneurs found that female entrepreneurs were evaluated as having less leadership ability and received less capital investment than similar male entrepreneurs. Moreover, studies suggest that female engineers who acted feminine were seen as incompetent [261, 346], whereas those who behaved in stereotypically masculine ways faced backlash [346, 288, 294, 77]. In addition, women's success at work was attributed to luck as opposed to skill [346].

Gender differences in employee evaluations at technology and business firms also indicated gender differences in perceived competency [294, 77]. While men received more actionable and task-oriented feedback, women received more critical and personality-related feedback. These findings were consistent across various industries, including military, politics, law, sports, and medicine [91, 291, 176, 47, 288, 38, 233]. The studies discussed above analyze either numeric (ratings), categorical (tags chosen from a predefined list of attributes), or textual performance reviews. Researchers who analyzed written performance reviews read the comments and coded them according to various parameters, including tone, valence, and skills discussed (technical, communal, agentic, and others). A drawback of these studies is that they are based on small datasets, with under 300 reviews of engineering professionals' postgraduate employment.

As seen above, men and women are perceived differently during engineering programs and careers. Research shows that reiteration of gendered feedback affects women's STEM identity, increases self-doubt, lowers self-efficacy beliefs, reduces performance, and thus, increases attrition [222, 288, 157, 90, 240, 222, 113, 108]. Since women are more likely than men to internalize the feedback they receive [239, 222], gendered evaluation may trigger stereotype threat in women, reducing their STEM self-efficacy beliefs and forcing them to confirm to these evaluations [302]. Furthermore, these low self-efficacy beliefs are exasperated by "weed-out" courses during engineering programs and everyday sexism in informal interactions with colleagues [310, 113, 285, 90]. Gender difference in perceived competency not only reduces women's self-confidence, performance, and in turn, willingness to persist in engineering programs and careers, but it may also reduce their opportunities to be hired or promoted, leading to dissatisfaction and attrition [288, 291, 157, 90]. Studies attribute these widespread gender differences in perceived competency and evaluation to unconscious yet pervasive pro-male STEM stereotypes and the under-representation of women in these fields [148, 200, 106, 107, 137]. Both positive and negative gender stereotypes have been shown to change students' self-image and lead them towards careers within or away from STEM [148, 339, 85, 222, 239].

2.2.2 Opportunity

Not only are STEM women perceived to be less competent than men, but they also receive fewer opportunities. Historical analyses attribute gender difference in opportunity to the masculine identity of the computing field as well as the cultural association of men with money and success [308, 347, 106, 107, 148, 200, 135]. Fewer opportunities, in terms of hiring, promotions, wages, and support (discussed below) prompt more women to leave engineering programs and careers.

In terms of hiring, there are conflicting reports on differences in the number and kind of job opportunities received by men and women. While some studies indicate that hiring practices, especially in academia, favour women [34, 349, 49], most labour market studies show a bias towards hiring men [232, 269, 303, 226, 190]. Researchers found that, based on just looking at a candidate, both male and female subjects (191 in total) were twice more likely to hire a man than a woman for a technical position [269]. A similar result was seen in academic hiring where both male and female subjects were more likely to vote to hire a male job applicant than a female with an identical record [303, 232]. Moreover, an experiment involving over 6,500 professors suggested that male faculty were (statistically significantly) more likely to ignore emails from prospective female graduate students [226]. Experiments to understand how venture capitalists evaluated entrepreneurial ventures revealed that female-led ventures were penalized relative to male-led ventures, in terms of lesser capital investment, as a result of a perceived lack of fit between female stereotypes and the expected personal qualities of entrepreneurs [190]. In addition, some qualitative studies found that women in engineering were relegated to managerial or secretarial roles more often than men [285].

Similar conflicting evidence was found in terms of gender differences in promotions in STEM jobs. While many studies find no gender differences in promotion probability [168, 255, 262, 27, 171], and still others conclude that the likelihood of promotion is higher for women [150, 300], the most common finding is that men are more likely to be promoted than women [43, 246, 223, 350, 258, 70, 266, 257, 23, 17]. Researchers speculate that discrepancies in the reported findings could be due to the lack of control variables, including industry, starting position, stage of career, voluntary job switches, career progression, and others [174]. Taking into account the aforementioned variables in an ordered logit model built on career path information of more than 20,000 employees of a company between 1981-2006, a study concludes that not only are men more likely to be promoted in their early careers [174], but they also have a higher chance of being promoted in their later undertakings [174, 257, 103, 257]. A focus group of 10 female faculty members suggested that the main cause for their reluctance in applying for, and in turn, delaying promotion, was the lack of feedback and mentoring [126]. Unlike their male peers, women managers' ability to mentor well did not increase their chances of promotions [103].

Moving on to difference in opportunity in terms of salary, the gender inequality in salary was undisputed, with men receiving higher wages than women in both STEM and non-STEM fields [15, 156, 232, 17, 170]. As suggested by the salary averages of 150-300 survey respondents, not only did female STEM professors receive lower starting salaries than male professors (even with equal likelihood of negotiation [251]), but also their salaries consistently remained lower in the later stages of their careers [232, 17]. While all studies agree with the gender gap in wage, the source is disputed. Some studies suggest that perpetuation of the gender gap in wages is unintentional and stems from an automatic stereotype that links men, more than women, with wealth [347]. Other quantitative studies that build statistical models on salary data (collected from the employees of a company, national surveys, or the census) use various control variables and suggest that these variables of career progression, including gender differences in starting salaries, voluntary job switches, career interruptions, fewer weekly working hours due to family reasons, and promotions, especially during early careers, lead to the gender wage gap [174, 223, 199, 211, 187, 175, 18, 354, 130, 234, 115, 43, 246, 350, 258, 70, 266, 257, 23]. Furthermore, men's propensity to search for and change jobs while being employed was found to be another reason for their increase in wages [174, 175]. Whatever be the source, regression analysis on salary information collected by a longitudinal survey suggests that as much as 75% of the gender wage gap remains unexplained by both academic and labour market variables [170].

In addition to unavailability of jobs and dissatisfaction over pay and promotion opportunities (discussed above), studies found working conditions to play a significant role in women's attrition from STEM careers. Both qualitative and quantitative studies, including those that analyze career paths of STEM graduates and the reasons behind their job transitions, confirm this finding [157, 132]. STEM workplaces are often lacking in support for women with young children and other care-taking responsibilities, which forces them to vacate STEM positions at greater rates than men [339, 317]. Since STEM fields are rapidly changing and require a substantial time commitment and continuous development of expertise to remain both productive and competitive [202], it is difficult for women to take maternity leave and maintain the productivity levels of their male and childless female peers [48]. In addition, female STEM workers report receiving less mentoring and support, which in turn increases gender differences in opportunities received [126, 17, 232]. The lack of female professors and colleagues initiates a feedback loop, which leads to fewer women enrolling in engineering programs and more women leaving engineering careers [45, 51, 58, 94, 111, 193, 297].

2.2.3 Satisfaction

In addition to the dissatisfaction caused by gender differences in perceived competency and opportunities received (discussed in Sections 2.2.2 and 2.2.2), studies show that women's persistence in engineering is negatively affected by its masculine environment and gender-based discrimination. Women who leave engineering programs cite several reasons for their dissatisfaction. These include the male-centric science pedagogy, the language surrounding STEM courses, lack of meaningful coursework or pathways to meaningful careers that can help others, the masculinity of orientation programs, the hostile and competitive atmosphere, lack of female peers and faculty members, a reduced sense of compatibility and support, and implicit and explicit sexism from professors and peers [69, 189, 137, 285, 158, 272, 90, 212, 254]. Consequently, women were more satisfied with engineering programs if they had a robust community where they could engage with peers to discuss course content, participate in undergraduate research programs, work towards their altruistic ambitions, and join STEM-related organizations, including single-sex programs [108, 272]. Interviews with 60 students noted that perceptions of being respected by course instructors positively influenced female students' intent to continue in engineering studies and even careers [4]. In addition, some women reported that they stayed in engineering programs for the prestige, creativity, and enjoyment of learning how to apply math and science to everyday life [290, 221]. Nonetheless, analysis of 1,629 survey responses revealed that satisfaction with the engineering major did not directly translate into pursuing a career in engineering, particularly among women [4].

Glass et al. [132] found women in STEM occupations to be significantly more likely to leave their occupational field than women in other fields, especially during early careers. Comparing the longitudinal career path information of 258 STEM and 842 non-STEM women graduates, the study found that while a similar proportion of women from both groups left jobs for family reasons, women from STEM left more often as job rewards, such as advanced training, failed to build commitment [132]. Additionally, studies have found that men and women feel differently about the content, pay, environment, and supervision provided at the job [218, 184]. In a meta-analysis of 31 studies, Konrad et al. [184] found that men considered earnings and responsibility to be more important than women did, whereas women considered prestige, challenge, task significance, variety, growth, job security, good co-workers, a good supervisor, and the physical work environment to be more important than men did. Thus, studies report that women who receive more workplace support, for example in terms of a supportive manager, are more satisfied and stay in engineering longer, indicating that satisfaction can affect retention [120, 10].

In addition to the dissatisfaction with masculine incentives, dissatisfaction with gender-based discrimination in pay and kind of opportunities received [157] and the masculine work and after-work culture [285] prompted women to leave engineering careers. Manually coding interviews and diary entries of 96 female engineering students suggested that they felt uncomfortable in male-dominated professions [285]. In particular, the qualitative analysis showed the effect of socialization during internships and how it lead women to develop less confidence about fitting into the culture of engineering [285]. In addition, the overt and implicit sexism, gendered expectations, and a lack of professionalism in the STEM workplaces prompt women to leave engineering careers [285, 293, 126].

2.2.4 Choice

More women than men *choose* to leave engineering programs and careers due to changes in their professional interest, career goals, and lifestyle values. A common reason stated by women leaving engineering programs is their discovery of an aptitude for a non-STEM major that seems better suited for their interests, talents, personality, educational, career, and life goals [158, 113]. Hunt [157], who conducted statistical analysis on over 200,000 observations collected via a national survey of STEM graduates, observed that changes in professional interests played a significant role in female attrition from STEM, even during later careers. In addition, women found engineering careers to be misaligned with their personal goals and values. While women in academia chose to publish less [138], women entrepreneurs expressed more caution than their male counterparts [238, 282]. Changes in women's career paths were also motivated by altruism and their preference to work with other people [337, 308]. For instance, a study that combined survey with census data found that women with a STEM degree were less likely than their male counterparts to work in a STEM occupation and were more likely to work in education or healthcare [15]. Nevertheless, we found no study indicating preference towards particular fields within STEM.

Additionally, some studies found women's lifestyle values of rearing a family and work-life balance to play a role in their choices [317, 118, 339, 157, 132, 101, 142]. Not only were women more willing than men to make occupational sacrifices for the sake of their families, but they also preferred less work-centred lifestyles [101, 142, 348]. Reports examining academic careers in STEM found more women to vacate tenure-track positions, and instead, opt for flexible part-time lecturer positions [178, 348]. Not only did more women than men graduates view STEM careers as unsuited to achieving their familial goals, but this difference increased as individuals entered their mid-30s or planned to have children [219, 118]. These gender differences were observed regardless of interest or performance.

Summary: As seen above, qualitative and quantitative studies based on surveys, interviews, experiments, organization-wide employee data, longitudinal surveys, and census data suggest that more women than men leave engineering programs and careers due to gender differences in perceived competency, opportunity, satisfaction, and choice. Not only do men receive better feedback from their professors and employers, but they also receive more opportunities in terms of job offers, promotions, and wages. While some women leave engineering due to their dissatisfaction with its masculine environment, incentives, and gender-based discrimination, others choose to leave because of changes in professional interest, people-oriented career goals, and lifestyle values. Each of the above-stated gender differences contributes towards the gender gap in engineering careers.

While most of the work listed above is based on engineering professionals' post-graduate employment and later career stages, studies that analyze career trajectories of STEM graduates, *(a)* note high rates of female attrition during early careers [132, 243], and *(b)* emphasize the importance of early career experiences in driving attrition [174, 157]. To fill this gap, we investigate gender differences in early engineering careers and examine whether the gender differences in opportunity, choice, perceived competency, and satisfaction - as have been observed in later careers - also exist during early careers. Since co-op work terms correspond to students' first experiences in the engineering workplace, this thesis focuses on co-op experiences to identify gender differences in early careers.

Co-op, as explained before, is a form of work-integrated learning based on the concept of "learning by doing" [183]. In co-op programs, students alternate between on-campus classes and off-campus work terms. Prior work has examined the benefits of co-op education from the perspective of three main stakeholders: employers, educational institutions, and students [318, 140]. For employers, co-op programs serve as a talent pipeline and a recruiting tool [140, 59, 61, 322, 98]. Research focuses on studying employers' expectations and aligning graduate competencies with employer needs [71, 61, 154]. From the institution's perspective, co-op programs serve as a tool for outreach and recruitment [140, 5, 44, 98, 265, 217, 289]. Thus, in order to understand why students apply to co-op programs in the first place, researchers survey existing students who are requested to think back to the time before starting their program and answer such questions retrospectively [274, 229, 5, 265, 217, 306, 95, 265]. Additionally, co-op programs help institutions enhance relationships with industry and align curricula to job market needs [140, 61].

Finally, from the students' perspective, co-op programs and their effects on student learning and career growth have been studied extensively. Co-op allows students to apply the concepts learnt in class to the real world and gain new technical, workplace, and job-seeking skills [263, 318, 3, 140, 196, 242]. Participation in co-op makes it easier for co-op students to find jobs after graduation, and that too, with higher salaries than non-co-op students [22, 127, 352]. While most research focuses on the impact of co-op on students' grades, skills, career planning, career growth, satisfaction, and retention in post-secondary programs [98, 318, 263, 3, 306, 127, 22, 117, 164], some studies investigate competition related to interviewing for and securing co-op placements and the stress it causes students [166, 322, 167, 253, 92]. While research has been conducted to enhance the effectiveness of co-op and improve co-op programs for all participating students [129, 147, 264, 163, 215], co-op has never been studied from a gender perspective.

This thesis is the first to study co-op from a gender perspective. Analyzing data from engineering students' co-op experiences, it is the first to investigate early engineering careers to understand gender differences in opportunity, satisfaction, perceived competency, and choice. In addition, it is also the first to study a (closed) labour market system containing information about *all* competing job candidates. While workforce literature includes qualitative studies that survey few job seekers to understand their decision-making during applications or interviews [14, 275], in this study, we follow all job candidates in a market as they move along all the stages of a job search process, including applications, interviews, subsequent shortlist, job offer, and offer acceptance. Thus, it is the first study to analyze gender differences in internal employment decisions and outcomes.

Chapter 3

Data

Our analysis uses three datasets. We have obtained two datasets from a large undergraduate North American institution located in Ontario, Canada, and one dataset from publicly available information. This section introduces each dataset and explains it in detail. Each dataset corresponds to a component of the educational pipeline, as shown in Figure 1.2. We combine the datasets to get a holistic view of the leaks in the pipeline.

1. **Admissions:** The dataset contains 33,763 records of admission forms submitted to the engineering faculty of the university (both accepted and rejected), from 2013 to 2016 inclusive. Each record contains the applicant's program of choice among the 13 available engineering programs, and their academic, extracurricular, and employment background. The details are discussed in Section 3.1.

2. **High School:** Made available through two online sources, the dataset contains public information on all public high schools in Ontario. It contains the school's aggregate performance on standardized tests, the gender gap in performance, and demographic information about its students. In conjunction with the Admissions dataset, this dataset provides additional background information regarding the engineering applicants. Further details are listed in Section 3.2.

3. **Co-operative education (Co-op):** This dataset is generated by the various processes of the undergraduate co-op pipeline of the institution, as outlined in Figure 1.2. Spanning over three semesters from September 2015 to August 2016 and ranging over 13 engineering programs with mandatory co-op, the dataset contains information regarding 8,956 students applying to 10,387 jobs. Other details of the dataset are discussed in Section 3.3.

3.1 Admissions Data

For each application, the Admissions dataset includes the gender of the applicant (male or female), their top choice of engineering program (among the 13 engineering programs offered at the institution, namely, Mechanical, Mechatronics, Computer, Electrical, Software, Nanotechnology, Geological, Systems Design, Civil, Chemical, Management, Biomedical, and Environmental), the name of the high school they attended, and short free text responses to the following fields:

1. Engineering interests and goals: Explain why you are interested in engineering and the specific program you applied to.
2. Reasons to apply to the university: Tell us about your reasons for applying to this university.
3. Programming experience: List any programming experience you have.
4. Extracurricular activities: List any extracurricular activities or areas of significant interest.
5. Jobs: List any jobs you held throughout high school.
6. Reading interests: Discuss a book or an article you enjoyed or that has had an impact on you (preferably something that was not part of a course at school).
7. Additional information: Tell us anything else about yourself that you would like us to know when reviewing your application.

3.2 High School Data

This dataset is obtained from two sources that host public information on all the public high schools in Ontario: (1) The Fraser Institute[1], a non-profit that studies and publishes the effects of new government policies and initiatives. As part of its initiative, it ranks various public secondary schools in Ontario, which along with other performance metrics of the schools, it makes publicly available[2]. (2) The Ontario Ministry of Education, that releases the performance and demographic data of all public high schools in Ontario[3].

[1] https://www.fraserinstitute.org/

[2] https://www.compareschoolrankings.org/

[3] https://www.app.edu.gov.on.ca/eng/sift/PCsearchSec.asp

Combining data from both these public sources gives us the following information about the 747 public high schools in Ontario. For each school, the High School dataset includes:

1. School name
2. Address (used to calculate the shortest driving distance from the school to the university).
3. Number of Grade 10 students who were *eligible* to write the province-wide Ontario Secondary School Literacy Test (OSSLT)[4] in 2016. The OSSLT tests English reading and writing skills.
4. Demographics (calculated using the 2016 Census):
   (a) Percentage of students who live in lower-income households.
   (b) Percentage of students whose parents have some university education.
   (c) Percentage of students whose first language is not English.
   (d) Percentage of students who are new to Canada (have moved to Canada within the last four years).
5. Academic performance:
   (a) Percentage of Grade 10 students who passed the OSSLT in their first attempt in 2014, and the gender gap in the percentage of students who passed.
   (b) Average score achieved by Grade 9 students in the province-wide Academic Math and Applied Math exams[5] written in 2013, and the gender gap in the scores.

To obtain the high school background of students who apply to the engineering faculty of the university from the public high schools in Ontario, we use the *name of the high schools* to merge the High School dataset with the Admissions dataset (Section 3.1). Between 2013 and 2016, the university received 17,814 applications from 670 public high schools in the province of Ontario.

For each high school in the merged dataset, we calculated the following metrics:

[4] https://www.eqao.com/the-assessments/osslt/

[5] https://www.eqao.com/the-assessments/grade-9-math/

1. *Proportion of engineering applicants*: Proportion of students in the graduating class who applied to an engineering program at the university. Since the High School dataset does not contain the size of the graduating class, we used the number of Grade 10 students eligible to write the OSSLT in 2016 as a proxy.

2. *Proportion of female engineering applicants*: Proportion of female engineering applicants among all engineering applicants from a school.

To reduce noise, we remove schools with fewer than 25 applicants per year, or fewer than 100 applicants in total over the four years.

3.3 Co-op Data

The Co-op dataset provides data regarding all the stages of the co-op pipeline at the university (shown in Figure 1.2). The co-op process and the data it generates are described below. Initially, employers participating in the co-op process submit job descriptions to the university, and any student (enrolled in a co-op program) may apply for any job. Next, employers select students they wish to interview. After conducting all interviews, employers indicate their preference for particular students by ranking them. A rank of zero, referred to as a "No Rank", means that the employer is not willing to hire the student. A rank of one, referred to as an "Offer", indicates that the employer wishes to hire the student. Ranks two to nine, referred to as "Ranks", represent the employer's backup or shortlist options, in order of preference. In other words, the employer would consider hiring these students if the top-ranked student declines the offer. Ranks do not need to be distinct, for example, an employer may put five students on the backup list and give all of them a rank of two.

After employers have submitted their rankings, the following information becomes visible to students. For each interview they participated in, a student is shown whether the corresponding employer made them an offer, shortlisted them (Rank - but the rank number is not shown), or is not willing to hire them (No Rank). Students then rank employers that made them offers or shortlisted them, between one and nine. As was the case with employer rankings, student ranks do not need to be distinct; for example, a student may give a rank of one to all of their options. In general, students assign a rank of one to jobs they are interested in, a rank of nine to jobs they strongly do not want, and ranks between two and eight to indicate an order of preference for the remaining options. An automatic student rank of zero is assigned to the No Ranks received.

The university then follows a matching process to assign students to jobs based on the rankings. The objective of the algorithm is to minimize the sum of the ranks of the resulting student-job assignment as a way to take the preferences of both parties into account. Students and employers are aware of this matching algorithm and thus would have ranked each other accordingly. Note that the lowest possible sum of ranks is two, and occurs when an employer offers a job to a student and the student gives a rank of one to this job. In this case, the student is guaranteed to be matched with this job[6]. Thus, the algorithm first considers rank pairs with a sum of two and matches these employers and students. Ties are broken randomly. Then, the algorithm considers the remaining unmatched jobs and students, scans for the lowest sums of ranks of three, four, and so on, and matches employers and students in that order. Consequently, students or employers may be matched with their first, second, or lower choice, or may not be matched at all.

Students work at the employers they were matched with for a four-month term. At the end of the work term, students and employers evaluate each other. Employers rate students' performance on various criteria and provide comments for future development. Students submit their satisfaction scores with the employer to the university confidentially. The Co-op dataset contains information regarding all the above stages of the co-op process in the schema shown below:

- Student data: anonymized student id, gender (male or female), program (among the 13 engineering programs mentioned in Section 3.1), number of work terms completed at application time (from 0 to 5)
- Job data: job id, job title, job description
- Application data: student id, job id
- Interview and ranking data: student id, job id, rank the employer gave to the student, rank the student gave to the job
- Placement data: student id, job id
- Employers' evaluations of students: student id, job id, numeric evaluation, and supervisor's comments

  Numeric evaluation includes:

[6] If a student were to give a rank of one to multiple Offers, the algorithm would randomly select one of these Offers.

  - an overall evaluation (on a 7-point scale: Unsatisfactory, Marginal, Satisfactory, Good, Very Good, Excellent and Outstanding Performance),
  - a detailed evaluation on the 16 criteria listed in Table 3.1 (on a 7-point scale, grouped into Developing (1-2), Good (3-5) and Superior (6-7), or "N/A" indicating not applicable).

  Text evaluation includes short free-text responses to the following questions:

  - Feedback: Please comment on the student's overall job performance in terms of their behavioral and developmental performance and expectations with respect to output, quality standards, delivery of goals and assignments.
  - Recommendations: Please provide your recommendations for the student's personal and professional development (optional).

  The evaluator's gender is unknown.

- Students' evaluations of employers: student id, job id, overall evaluation from 1-10 (10 being most satisfied).

Additionally, we have two semesters of data from a pilot program run from January to August 2017 to explore students' satisfaction with their work terms. Let us call this the *2017 satisfaction dataset*. It consists of 9,800 students with information about their gender, academic program, and seniority, but not their applications, interviews, or performance evaluations. However, in addition to an overall satisfaction score, it contains students' satisfaction scores for the following questions (rated between 1 to 5, with 5 being most satisfied):

1. Availability of employer support
2. Opportunities to learn or develop new skills
3. Opportunities to make meaningful contributions at work
4. Opportunities to expand your professional network
5. Appropriate compensation and/or benefits
6. How closely was your work related to your academic program
7. How closely was your work related to the skills you are developing at university

While there is an overlap of students between the *Co-op* and the *2017 satisfaction* datasets, some students from the Co-op dataset have graduated by 2017, and there are new students who enrolled in Fall 2016 and had their first work terms in 2017. We use the 2017 satisfaction dataset in conjunction with the Co-op dataset to provide additional insights on gender differences in satisfaction with co-op.

Before using the data, it needed to be cleaned extensively. Particularly, we noted some data quality issues in the Co-op dataset. These included, missing values in records, missing records for some students, duplicate records, stale or erroneous information in manually entered data, badly formatted text data, and use of proxies/flags to accommodate for unforeseen business needs.

We use the Co-op dataset to measure gender differences in the co-op experiences of the entire student population (i.e., all engineering students or ENG) as well as specific sub-populations. Since we are analyzing students' work experiences, the institution provided us with a mapping that grouped similar academic programs. Similar academic programs are those whose students compete for the same set of co-op jobs and eventually work in the same industry. From hereon, these grouped academic programs are referred to as disciplines. The 13 academic programs offered by the institution were grouped into nine job disciplines; students who were enrolled in Computer and Software Engineering programs were classified into one job discipline labelled COMP, and students who were enrolled in Mechanical and Mechatronics Engineering programs were classified into another job discipline labelled MECH. In addition, students from the academic programs of Systems Design and Management were grouped into a discipline labelled Industrial and students from the Environmental and Geological academic programs were grouped into Environment.

In addition, seniority is measured in terms of the number of work terms completed rather than the academic level: junior students are those who have completed 0 or 1 work terms and senior students are those who have completed at least 4 work terms. The sizes and the gender mix of the different sub-populations within ENG are summarized in Tables 6.1 and 6.2 of Section 6.

3.4 Limitations

The nature of our data introduces limitations to our work:

- The institutions that collected the datasets used in our study recorded gender as a binary male or female attribute. We recognize that gender is not binary and is not interchangeable with sex. However, as a result of this data limitation, we only report results for male and female students.

- Due to the social construction of gender, it is greatly impacted by society, culture, and personal experiences. Thus, the inferences drawn from these datasets might not apply to other countries or cultures.

- The datasets studied in this thesis correspond to a single North American university with a large engineering program. In addition, participation in co-op is mandatory for all engineering students at the university. Therefore, inferences drawn from this study may not apply to other disciplines, co-op programs, or institutions.

- The data extracts used in this study did not include certain attributes, for example, the students' grade point average (GPA), the co-op supervisor's gender, or the extent of mentorship, making it impossible to control for related factors that could have affected the outcome. Similarly, the proportion of female applicants from a high school may have been affected by factors other than those observed in the data, for example, the number of women STEM teachers in the school.

- The datasets do not provide information on certain fields, thus, forcing us to use proxies. For example, a job posting just mentions the job title and description and does not contain a structured input for discipline. Thus, to analyze gender differences in particular disciplines, missing discipline labels forces us to come up with our own definitions.

- We assume that the job description provided by employers is representative of the job. However, the actual nature of the job and its requirements may be different. Similarly, we assume that responses provided by applicants or supervisors are representative of their thoughts, even though their actual reasons may be different.

- A general limitation of secondary data analysis is that they focus on the question of "what" rather than "why". In other words, they can identify interesting patterns and correlations in the data, but not cause-and-effect relationships. Our data-driven analysis has similar limitations where we identify frequent patterns in the comments and experiences of male and female students, but can only speculate on the reasons behind them. For example, we do not exactly know why a student applied to a particular job or why they evaluated it poorly. While our data-driven analysis provides a starting point for further study, only interviews, which are out of the scope of the study, can provide ground truth.

Table 3.1: Student performance evaluation criteria

Evaluation Criteria	Description
Interest in Work	The degree to which the student pursues goals with commitment and takes pride in accomplishments
Ability to Learn	The extent to which the student becomes proficient with job duties and work processes
Quality of Work	The ability of the student to set high standards for own personal performance; strive for quality work; put forth extra effort to ensure quality of work
Quantity of Work	The volume of work produced by the student, along with his or her speed and consistency of output
Problem Solving	The student's demonstrated ability to analyze problems or procedures, evaluate alternatives, and select the best course of action
Teamwork	The degree to which the student works well in a team setting
Dependability	The manner in which the student conducts his or herself in the working environment
Response to Supervision	The manner in which the student responds to direction and constructive criticism
Reflection	The student's demonstrated ability to learn and adapt from previous mistakes and experiences
Resourcefulness	The student's demonstrated ability to develop innovative solutions and display flexibility in unique or demanding circumstances
Ethical Behaviour	The extent to which the student's behaviour demonstrates integrity and ethics in work and relationships
Appreciation of Diversity	The degree to which the student shows understanding and sensitivity to needs and differences of others (i.e., ethnicity, religion, language, etc.)
Entrepreneurial Orientation	The student's demonstrated ability to take informed risks and demonstrate creativity and add value to the company
Written Communication	The extent to which the student demonstrates effective written communication
Oral Communication	The extent to which the student demonstrates effective oral communication
Interpersonal Communication	The extent to which the student effectively listens, conveys, and receives ideas, information, and direction

Chapter 4

Methods

Our study uses standard statistical and text analysis tools. They are described in Sections 4.1 and 4.2 respectively.

4.1 Statistical Analysis

The statistical tools we use include the Pearson correlation coefficient, the two-proportion z-test, the t-test, and the Mann-Whitney test. For all the statistical tests, we use a p-value of 0.05 as a threshold to determine significance and asterisks to indicate their strength (*** suggests a p-value less than 0.001, ** suggests a p-value less than 0.01, and * suggests a p-value less than 0.05).

Correlation coefficient: The Pearson correlation coefficient measures the strength of linear association between two (continuous) variables and ranges between +1 to -1; zero indicates no association, positive values indicate a positive relationship, and negative values indicate a negative relationship (the larger the magnitude of the correlation coefficient, the stronger the relationship). In addition, a p-value less than 0.05 confirms that the correlation coefficient found is statistically significant. The null hypothesis of the significance test states that the correlation coefficient between the two variables is in fact zero.

We use this method to identify unique characteristics of schools that produce many female engineering applicants (Section 5.3). Consequently, we calculate the Pearson correlation coefficient between the percentage of female engineering applicants from each school and the school's aggregate demographic and performance metrics (available in the high school dataset described in Section 3.2).

Two-proportion z-test: A two-proportion z-test determines whether two proportions are different from each other. We use this test to examine gender differences in students' co-op experiences (Section 6), where we compare: (a) the fraction of male and female students who obtained at least one interview, (b) the proportion of female students who worked in a particular type of job with the proportion of female students in that discipline, (c) the proportion of male and female students who received a "N/A" score on particular evaluation criteria, and other fractions.

T-test: A t-test is used to compare the averages of two groups and determine if the difference is statistically significant. For example, we use a t-test to compare the average number of applications submitted by male and female students, the average number of top-3 ranks received by male and female students, and the average interview to offer conversion rate (i.e., the number of offers received by a student divided by the number of interviews they received). Other applications of t-tests can be found in Section 6.

Mann-Whitney test: The Mann-Whitney test is used to compare the average scores of two groups when the scores are measured on the Likert scale. Since all the evaluation scores in our dataset, including the employer's evaluation of the student (overall and the 16 specific performance criteria) and students' satisfaction with their employer (overall and the seven individual aspects) have an ordinal nature (shown in Section 3.3), we use the Mann-Whitney tests to compare the average scores given and received by male and female students.

The above statistical tools are used to measure gender differences in the entire population (i.e., all engineering students or ENG) as well as specific disciplines, including COMP and MECH since they are the largest disciplines in the datasets.

4.2 Text Analysis

We apply syntactic text analysis methods to all the free-text fields of our datasets. To that end, we analyze (a) *applicants' responses in admission forms* to identify gender differences in their motivations and interests as well as changes in these responses with the number of female engineering applicants from high schools, (b) *job titles and descriptions* to identify differences between co-op jobs that receive more applications from or employ more male and female students, and (c) *supervisor's comments* to identify gender differences in early career performance reviews. Since all these fields have a free-text format, we first convert each document (or response) into a set of standardized word forms (Section 4.2.2) and then examine the syntactic gender differences between different pairs of groups (Section 4.2.3). Section 4.2.1 describes the methods used to identify these groups.

4.2.1 Identify Groups

In addition to comparing the documents associated with all ENG male students and all ENG female students, we analyze gender differences in other groups. For example, when analyzing gender differences in applicants to engineering programs (Section 5), we group them by their program of choice and examine gender differences within each group separately. Similarly, when studying gender differences in students' co-op experiences (Section 6), we group students by their job disciplines and seniority (definitions available in Section 3.3). To avoid overfitting, we ensure that each group has more than 100 students. We present observations from groups only when they differ from the gender differences seen in the entire population (i.e., all of ENG).

Certain research questions require identification of other groups. For example, to identify the unique characteristics of schools that produce many female engineering applicants based on what applicants from these schools said in their applications (Section 5.3), we first need to identify schools that produced a much higher or lower proportion of female engineering applicants. To do this, we consider the 31 schools that produce at least 100 engineering applicants between 2013 and 2016. We sort these schools by the proportion of female engineering applicants they produced and calculate the 25th percentile (i.e., the lower quartile or Q1) and the 75th percentile (i.e., upper quartile or Q3). Schools with a proportion of female engineering applicants lower than Q1 were labelled M_schools, and schools with a proportion of female engineering applicants greater than Q3 were labelled F_schools. Both groups, namely M_schools and F_schools, contained eight schools and more than 1000 applicants each. By definition, there are more male applicants in M_schools and more female applicants in F_schools. Therefore, to identify the unique characteristics of F_schools (i.e., the schools that produced many female engineering applicants), we compared the responses of male applicants from M_schools to the responses of male applicants from F_schools. Separately, we compared the responses of female applicants from M_schools and F_schools.

Similarly, to understand if male and female students apply to different kinds of co-op jobs, we first need to identify the groups of co-op job postings that received a much higher proportion of applications from male students versus female students (Section 6.1). We conduct this analysis for COMP and MECH since they are the two largest disciplines in the dataset. To begin, we identify job postings that belong to particular disciplines (recall, the job postings in the dataset do not include industry or discipline labels). Discipline labels are created as follows: If a job posting received at least 10 applications from (junior or senior) students of a particular discipline, this job was said to belong to the corresponding (junior or senior) discipline. For example, a senior COMP job must have at least 10 senior

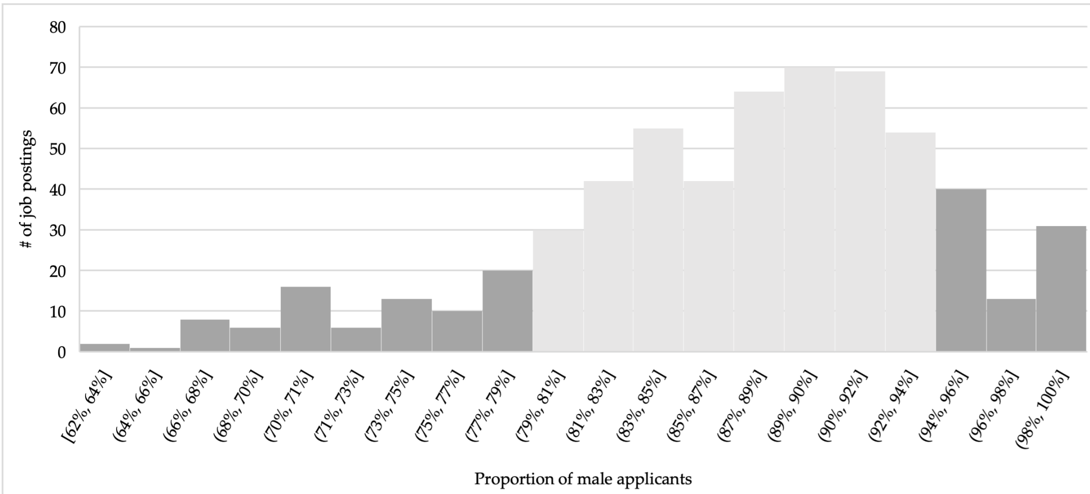

Figure 4.1: Distribution of the proportion of male applicants in senior COMP jobs

COMP students applying to it. Even though this method is not perfect (e.g., a project management job that received 10 applications from COMP students would be considered a COMP job), other labelling methods that were tested were even less precise. For example, using the discipline of the student who obtained a job is sensitive to outliers: a MECH student may have obtained a software developer job that mostly COMP students applied to. Likewise, relying on the presence of particular keywords was problematic due to the lack of an exhaustive list of discipline specific skills. We use the method explained above to identify the job postings that belonged to the following disciplines: COMP (containing 3232 job postings), junior COMP (2267), senior COMP (592), MECH (1657), junior MECH (912), and senior MECH (395). While the junior and senior jobs of a discipline are strict subsets of that discipline, there are a few jobs that appear in both.

After identifying the job postings that belong to a particular discipline, we identify job postings that receive a much higher proportion of applications from male students (referred to as jobs_{M}) and female students (referred to as jobs_{F}) of that discipline. For each discipline, the distribution of the *proportion of male applicants to every job posting* is visually inspected to identify where the distribution function dropped off. For example, Figure 4.1 shows this distribution for senior COMP jobs, with the bulk of these jobs receiving between 79 and 94% of applications from male students; the distribution drops on either side of this range, suggesting the thresholds for jobs_{M} and jobs_{F}. Additionally, to avoid overfitting, we ensure that jobs_{M} and jobs_{F} have more than 50 job postings.

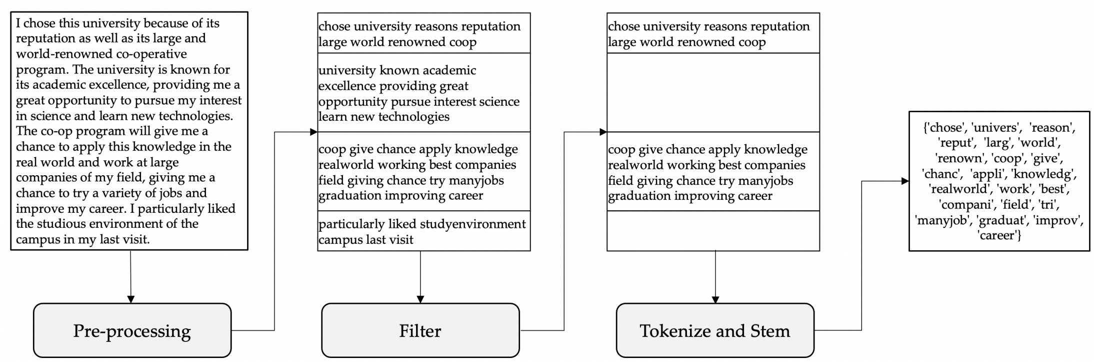

Figure 4.2: Parser used to convert a free-text document into tokens

4.2.2 Convert free-text to tokens

Now that we have identified the groups to be compared, this section converts the documents of each group to tokens, and the next section identifies significant differences between them. We implemented a parser in Python to extract words from free-text fields. The parser uses standard text mining techniques [79] to convert each document into a set of standardized word forms (referred to as "words", "tokens", or "terms" in the remainder of the thesis). Figure 4.2 summarizes the three steps of the parser, described below, and shows the outcome of each step when the parser is applied to a particular document. The document shown in Figure 4.2 is a response to the question "Tell us about your reasons for applying to this university". To preserve data privacy, the figure shows a synthetic response similar in style to those in the dataset.

The operations performed in each of the three steps are as follows:

Step 1 - Pre-processing: For every non-blank document, the parser performs the following actions.

1. The text is converted to lowercase.

2. Stopwords, which are words that serve a grammatical purpose but do not contain any meaningful information, such as words including "and", "the" and "is", are removed. In addition, words common in the context of the study, including "work term", "university", and "apply" are also removed.

3. Various forms of certain words and phrases are converted to a common form using

regular expression matching[1] (e.g., occurrences of "inter-personal" and "inter personal" are converted to "interpersonal"; "mathematics", "maths", and "math" are converted to "math"; "ap", "a-p", "a.p.", and "advanced placement" are converted to "ap"; "co-operative education", "co-op", "cooperative program", and "cooperative" are converted to "coop"; and "web development", "website development", "web developer", and "web-development" are converted to "webdev").

In addition to using domain knowledge, such words and phrases were identified using a natural language processing technique called n-grams [79]. An n-gram is a sequence of n consecutive words in a sentence [79]. Hence, for every text field, we first computed the unigrams (or single words), bigrams (pairs of consecutive words in a sentence), and trigrams (three consecutive words in a sentence) in its documents, removed the ones with a frequency of less than 1%, and identified those with similar meaning. We then used regular expression matching to combine the bigrams and trigrams with similar meaning into a common form.

4. The text is broken into sentences. The symbols, including the period ("."), question mark ("?"), and exclamation mark ("!") are used to tokenize the text into separate sentences.

5. Finally, special characters, digits, and punctuation are replaced by white space in each sentence.

The above operations are performed for all text fields, except job titles and descriptions. Since job descriptions are written directly by employers, they are neither standardized nor well-structured, and need more pre-processing. In particular, job descriptions may include information that is unrelated to the nature of the job, such as links to company websites, contact names and emails, timestamps and addresses, annotations with administrative content, special characters used for formatting, common abbreviations, misspellings, and of course common English words. In our prior work on job description mining [61, 59], we developed a method to extract informative terms from job descriptions. These informative terms included, required technical skills, soft skills, perks (e.g., free food or proximity to public transit), and other terms indicating the nature of the job and company culture. Instead of the standard *pre-processing* described above, we use this method to extract relevant tokens from job titles and descriptions [59].

Step 2 - Filter: This step retains sentences containing a particular token. Let us consider the example shown in Figure 4.2. As can be seen in Figure 4.2, the *pre-processing step*

[1] https://docs.python.org/3/library/re.html

of the parser returns separate sentences, each of which describes why a student applies to the university. Since the research question shown in Figure 1.1 particularly focuses on investigating gender differences in applicants' views on *co-op*, this step filters and retains only those sentences that contain the token "coop" (note that the previous step would have already converted the alternative forms of the token to "coop"). This step of the parser is optional and is used only when sentences that most likely contain certain topics need to be extracted from documents.

Step 3 - Tokenize and Stem: Finally, the text is tokenized by white space and stemmed using the NLTK snowball stemmer[2] [260]. Stemming converts related words with common meanings but different endings to a common stem. For example, the words "efficient", "efficiently", and "efficiency" are converted to "effici"; "expect", "expected", and "expectation" are converted to "expect"; "learned", "learning", and "learnt" are converted to "learn"; and "studying", "studious", and "study" are converted to "studi". The snowball stemmer was used instead of other stemmers (for e.g. Porter and Krovetz) because it converted many more related words to a common form. To illustrate, the Porter stemmer did not reduce the word "learnt" to "learn" and the Krovetz stemmer did not reduce the word "prediction" to "predict".

Overall, the parser converts a document to a set of tokens. The next section describes how these parsed documents are used to find differences between groups.

4.2.3 Identify Differences between Groups

This section introduces two methods, namely term frequency analysis and document clustering, that are used to identify differences in the text associated with two groups. For every text field (for example, Engineering Interests and Goals and job descriptions), both methods use the output of the parser (a document converted to a set of tokens) as input and identify differences between the words associated with the groups. In general, we compare documents belonging to ENG male and ENG female students. However, Section 4.2.1 identifies other pairs of groups to be compared.

For each pair of groups being compared (for example, the Engineering Interests and Goals of male applicants from M_schools and F_schools, or job descriptions of jobs filled by male and female students), we conduct a *term frequency analysis* and report the following:

1. *Frequent terms:* sorted by their frequency, these are tokens that occurred at least once in a large percentage of documents in the field, and

[2] https://www.nltk.org/_modules/nltk/stem/snowball.html

2. *Significant differences:* these are tokens that were present more frequently in one group than another. A two-tailed two-proportion z-test (described in Section 4.1) is used to determine whether the difference in token frequencies is statistically significant. The null hypothesis of this test states that the proportion of documents that contains a particular token is the same in both groups. Only differences that were statistically significant at a p-value of 0.05 were reported. Differences that were statistically significant at a p-value of at least 0.05 were marked with *, at least 0.01 with **, and at least 0.001 with ***.

For each token of interest, we use manual inspection to understand the context of its usage. Manual inspection entails (a) examining all the bigrams and trigrams with more than 1% frequency that contain the token, and (b) reading a random sample of at least 50 sentences from the original documents of the text field that contain the token .

In addition to the *term frequency analysis*, we conduct *document clustering* on the job descriptions of filled co-op positions. To understand if male and female students of the nine disciplines work in different types of cop-op jobs, we cluster the job placements of each discipline and report differences in the proportions of male and female students in each job cluster. The *document clustering* process is described below:

1. Recall that the parsing process (explained in Section 4.2.2 above) converts a job description into a set of relevant tokens. We use this set of tokens to obtain a *job vector* for every job description. Each coordinate of the job vector corresponds to a particular token amongst all possible tokens of the job description corpus. The ith coordinate of a job vector is equal to the token's inverse document frequency (IDF), provided that this token is present in the given job description (and zero otherwise). IDF is a common metric in information retrieval and represents the popularity of the word in the corpus [79]. We use this metric so that the clustering process is able to distinguish between common (e.g., "software", "development") and unique (e.g., "python", "communic") job attributes.

2. Following previous work on text clustering [89, 296, 298], we use Latent Semantic Analysis (LSA) to reduce the dimensionality of the job vectors, from the number of distinct words in the job description corpus down to one hundred. Each reduced dimension corresponds to a latent concept in the job description corpus.

3. We then run k-means clustering on the transformed job vectors and report the top ten ranking terms from each cluster centroid as representatives. We vary the value of K from 3 to 15 clusters and manually choose the K with the most comprehensible clusters. Each cluster represents a type of job in the discipline's co-op job market.

4. We manually assign a label to each cluster. To do this, we inspect the top terms of the cluster centroid (as well as the frequent bigrams, trigrams, and sentences from the job postings of the cluster) and assign a label that most likely represents the jobs of the cluster. For example, terms such as "html", "javascript", "css", and "sql" are most frequently found in web programming, and thus, a cluster centroid containing these terms could indicate a cluster of Web Programming jobs. Similarly, terms such as "databas", "busi", "c#", and "sql" are most likely to correspond to jobs related to Business Systems.

5. Finally, for each cluster, we calculate its size and the percentage of female students in it. We use a two-tailed two-proportion z-test (described in Section 4.1) to compare the percentage of female students in each cluster with the percentage of female students in the job discipline. We report all differences and mark clusters with significantly more male and female students with asterisks.

In addition to the methods described above, we tried various others. For instance, as can be seen in our paper on gender differences in engineering applicants [60], in addition to the term frequency analysis, we used a topic modelling technique to identify gender differences in students' motivations to apply to engineering. Since both methods generated similar results, in this thesis, we chose to report results from the simpler, and thus, more interpretable method.

Chapter 5

Gender differences prior to post-secondary education

The gender gap in STEM education and workforce is well-documented (Section 2). Studies have shown that, despite having high grades in STEM subjects, fewer women apply to and obtain STEM degrees: only 23% of women with high mathematics scores in high school pursue STEM degrees compared to 45% of men with the same scores [259, 144]. To understand why this is the case, we analyze gender differences in high school backgrounds and interests of undergraduate engineering applicants.

Following Figures 1.1 and 1.2, this chapter focuses on identifying gender differences in secondary education by answering three research questions. *First*, in Section 5.1, we use the applicant responses in the Admissions dataset to determine whether male and female applicants identify different reasons for applying to an engineering program and whether they have different technical and extracurricular interests. Table 5.1 shows the number of applications and the gender distribution of applicants to each program (sorted by percentage of female applicants). *Second*, we examine the percentage of male and female applicants who mention co-op as a reason for applying to the university and inspect any gender differences in their expectations from these programs (Section 5.2). Since all the engineering programs at the university have a mandatory co-op program, gender differences in the perception of co-op could affect the gender proportion of applicants. *Third*, we merge the Admissions and High School datasets to identify the unique characteristics of schools that produce many female engineering applicants (Section 5.3). While the works presented in Sections 5.1 and 5.2 have been published [60, 67, 64], the analysis shown in Section 5.3 is currently under review in the Journal of Engineering Education.

Table 5.1: Gender breakdown by program

Program	Applicants	%Male	%Female
Mechanical	5473	88%	12%
Mechatronics	2886	88%	12%
Software	3635	86%	14%
Computer	3931	84%	16%
Electrical	3782	83%	17%
Nanotechnology	1670	76%	24%
Geological	361	75%	25%
Civil	3375	72%	28%
Management	1040	64%	36%
Chemical	3612	62%	38%
Systems Design	957	62%	38%
Biomedical	2015	48%	52%
Environmental	1021	47%	53%
Total	33758	77%	23%

Overall, the goal of this chapter is to measure gender differences in the motivations, interests, and backgrounds of applicants to engineering programs. Since most of the previous work on identifying gender differences in students' motivations and interests has either been qualitative or has used small datasets, applying text mining methods to a large admissions dataset can help us find data-driven evidence for known differences as well as obtain new insights. The observations may provide actionable insights into increasing the number of female applicants to engineering programs by (a) identifying characteristics of student backgrounds that promote engineering interest in female students, and (b) recognizing ways to align engineering co-op programs with the expectations of female students.

5.1 Gender differences in the applicants to co-operative engineering programs

5.1.1 Motivation

As mentioned above, undergraduate engineering programs receive applications from only 23% women. As a result, there has been a great deal of research on understanding why this is the case. Existing literature identified reasons including, the negative stereotypes about women's STEM ability, low perceptions about the social impact of STEM careers, lack of role models, and lack of confidence in background knowledge as key reasons behind the under-representation of women in STEM programs [39, 87, 99, 221, 290, 339, 85]. However, since most of these results are qualitative and based on small datasets collected through surveys and longitudinal studies, we study this topic by applying text mining methods to a large admissions dataset containing 33,763 applications. In this section, we compare the text in the admission forms of male and female applicants to understand if they differ in their motivations for joining engineering, their interests, and backgrounds.

5.1.2 Data and Methods

Our analysis is enabled by the Admissions dataset described in Section 3.1. It contains applicant responses to questions regarding their engineering interests and goals, reasons to apply to the university, programming experience, extracurricular activities, job experience, reading interests, and other interests. *First*, we follow the process shown in Figure 4.2 and described in Section 4.2.2[1] to convert each non-blank response of the seven text fields to a set of tokens. *Then*, we conduct a term frequency analysis for every text field and report the *Significant Differences* between the responses of (all) male and female applicants (process described in Section 4.2.3)[2]. Common English words with significant differences are excluded from the report for brevity.

5.1.3 Results

Table 5.2 shows some tokens with frequency differences in the responses of male and female applicants for each of the seven questions in the Admissions dataset. For example, Ta-

[1] The optional "Filter" step is not used for any of the seven free-text fields

[2] We initially computed the gender differences in applicants to each program separately, but since they followed similar trends as observed in all applicants, their results were omitted for brevity.

ble 5.2a shows tokens that are mentioned significantly more frequently by male applicants than by female applicants (on the left) and vice versa (on the right) when discussing their engineering interests and goals. The lists are sorted by the difference in frequencies, abbreviated Δ and computed as the percentage of male (or female) applicants who mention a token minus the percentage of female (or male) applicants who mention this token. For example, the first row in the table on the left in Table 5.2a shows that male applicants mention "robot" 7% more often than female applicants. Similarly, the table on the right shows that female applicants mention"peopl" (word stem of "people") 6% more often than male applicants. The asterisks beside the differences in frequencies indicate the strength of the statistical significance of the difference (recall that *** suggests a p-value of less than 0.001, ** suggests a p-value less than 0.01, and * suggests a p-value less than 0.05). Below, we discuss the word frequency differences in the seven text responses of male and female applicants.

Engineering Interests and Goals: When describing why they want to study engineering, both male and female applicants specify technical interests. As can be seen in Table 5.2a, male applicants are more likely to mention words including "robot", "machin", "build", "hardwar", "game", "code", and "function". Correspondingly, female applicants mention their "love" of science using words such as "science", "research", "love", "problemsolv", "lab", and "innov" (some of these words can be seen in Table 5.2a).

Female applicants report other reasons to study engineering, more often than male applicants. As can be seen in Table 5.2a, female applicants are more likely to mention non-technical words including "people", "creativ", "improve", "help", "parent", and "chang". In addition, they also mention words including "women", "explor", "encourag", "contribut", "impact", and "societi" more frequently than male applicants. Manual inspection of responses containing these words confirm that, in addition to technical interest, female applicants want to join engineering due to parental influences and their desire to contribute to society.

Furthermore, while male applicants refer to pursuing engineering as their "goal" or "childhood dream", female applicants refer to it as a "challenge" (see Table 5.2a). Similar gender differences in motivations to join engineering were found in applicants to individual programs (results omitted for brevity).

Reasons to apply to the university: Both male and female applicants mention technical interests as a reason to apply to the university (see Table 5.2b). Male applicants are more likely to mention words including, "hardwar", "machin", "robot", "code", and "game", and female applicants are more likely to mention words including "studi", "scienc", "love", "innov", "research", and "passion". In addition to technical words, female applicants are

Table 5.2: Differences in frequencies between tokens mentioned by male and female applicants

(a) Engineering Interest and Goals

Token	Male	Female	Δ
robot	16%	9%	7%***
goal	25%	20%	5%***
machin	11%	7%	4%***
build	25%	21%	4%***
hardwar	5%	2%	3%***
game	6%	3%	3%***
childhood	6%	4%	2%***
code	5%	3%	2%***
toy	4%	2%	2%***
function	8%	6%	2%***

Token	Female	Male	Δ
scienc	42%	34%	8%***
peopl	21%	15%	6%***
creativ	16%	11%	5%***
research	16%	11%	5%***
love	25%	20%	5%***
improv	16%	12%	4%***
help	28%	24%	4%***
challeng	17%	13%	4%***
parent	7%	4%	3%***
chang	9%	6%	3%***

(b) Reasons to apply to the university

Token	Male	Female	Δ
reput	26%	22%	4%***
hardwar	4%	1%	3%***
machin	6%	3%	3%***
robot	6%	3%	3%***
goal	44%	41%	3%***
world	36%	34%	2%***
childhood	7%	5%	2%***
industri	9%	7%	2%***
prestigi	5%	4%	1%***
recommend	2%	1%	1%***

Token	Female	Male	Δ
opportun	36%	29%	7%***
coop	62%	57%	5%***
studi	41%	36%	5%***
scienc	45%	40%	5%***
love	22%	17%	5%***
help	29%	25%	4%***
innov	21%	17%	4%***
peopl	16%	12%	4%***
allow	23%	19%	4%***
research	20%	17%	3%***

Table 5.2: Differences in frequencies between tokens mentioned by male and female applicants, continued

(c) Programming Experience

Token	Male	Female	Δ
java	55%	40%	15%***
languag	38%	27%	11%***
game	24%	15%	9%***
c++	25%	16%	9%***
develop	16%	8%	8%***
learn	36%	29%	7%***
python	23%	16%	7%***
project	20%	13%	7%***
c	14%	8%	6%***
android	7%	2%	5%***

Token	Female	Male	Δ
mark	30%	25%	5%*
help	8%	6%	2%*
(coding workshop)	4%	2%	2%*
programmingconcept	5%	3%	2%*
attend	2%	1%	1%*
editingsoftwar	1%	0%	1%*
love	1%	0%	1%*
webdevelop	1%	0%	1%**
(high school CS course)	1%	0%	1%*
encourag	1%	0%	1%*

(d) Extracurricular activities

Token	Male	Female	Δ
robot	12%	6%	6%***
comput	8%	3%	5%***
game	5%	3%	2%***
coach	5%	3%	2%***
team	31%	29%	2%**
band	13%	11%	2%**
develop	2%	1%	1%***
intramur	2%	1%	1%***
video	2%	1%	1%***
websit	2%	1%	1%***

Token	Female	Male	Δ
art	15%	6%	9%***
council	20%	12%	8%***
danc	10%	2%	8%***
music	16%	9%	7%***
communiti	21%	15%	6%***
volunt	23%	17%	6%***
fundrais	14%	8%	6%***
lead	19%	13%	6%***
children	7%	3%	4%***
editor	4%	1%	3%***

Table 5.2: Differences in frequencies between tokens mentioned by male and female applicants, continued

(e) Jobs

Token	Male	Female	Δ
comput	5%	2%	3%***
refere	4%	2%	2%***
labour	3%	1%	2%***
repair	3%	1%	2%***
mainten	4%	2%	2%***
technician	3%	1%	2%***
grocer	5%	3%	2%***
waiter	2%	0%	2%***
cook	2%	1%	1%***
landscap	1%	0%	1%***

Token	Female	Male	Δ
server	8%	1%	7%***
cashier	13%	7%	6%***
assist	18%	14%	4%***
teacher	6%	3%	3%***
babysitt	3%	0%	3%***
counsellor	4%	2%	2%***
camp	6%	4%	2%***
research	2%	1%	1%***
receptionist	2%	1%	1%***
swim	4%	3%	1%**

(f) Reading Interests

Token	Male	Female	Δ
articl	18%	13%	5%***
comput	6%	2%	4%***
enjoy	29%	26%	3%***
scienc	12%	10%	2%***
physic	7%	5%	2%***
explain	6%	4%	2%***
creat	12%	10%	2%***
design	5%	3%	2%***
space	4%	2%	2%***
theori	4%	3%	1%***

Token	Female	Male	Δ
love	21%	13%	8%***
novel	31%	25%	6%***
charact	20%	15%	5%***
women	6%	1%	5%***
stori	31%	26%	5%***
peopl	29%	25%	4%***
famili	11%	7%	4%***
perspect	10%	7%	3%***
societi	15%	12%	3%***
emot	6%	4%	2%***

Table 5.2: Differences in frequencies between tokens mentioned by male and female applicants, continued

(g) Additional Information

Token	Male	Female	Δ
comput	11%	6%	5%***
sport	10%	6%	4%***
team	16%	13%	3%***
robot	5%	2%	3%***
game	5%	2%	3%***
physic	11%	9%	2%**
code	2%	1%	1%***
video	2%	1%	1%***
solv	4%	3%	1%***
websit	2%	1%	1%***

Token	Female	Male	Δ
art	7%	3%	4%***
volunt	11%	7%	4%***
passion	14%	11%	3%***
love	14%	11%	3%***
peopl	16%	13%	3%***
learn	30%	27%	3%***
danc	3%	1%	2%***
communiti	11%	9%	2%***
famili	10%	8%	2%***
hardwork	12%	10%	2%***

more likely to mention tokens that indicate altruistic tendencies (for example, “help” and “peopl” in Table 5.2b).

In addition, male applicants mention the reputation of the university as a reason to apply, more often than female applicants. This is indicated by words including, “reput”, “world”, “industri”, “prestigi”, “recommend”, “childhood”, “technolog”, and “friend” (some of these words can be seen in the table on the left in Table 5.2b). In addition, male applicants mention the university’s start-up culture more often (indicated by words including “compani”, “entrepreneur”, and “startup”).

On the other hand, female applicants mention the university’s co-op program and its benefits, more often than male applicants (see table on the right in Table 5.2b). Words including, “coop”, “opportun”, “allow”, “practic”, “explor”, and “financ” occur more frequently in the responses of female applicants (some of these words can be seen in the table on the right in Table 5.2b). Gender differences in applicants’ perceptions of co-op are explored further in Section 5.2.

Programming Experience: As can be seen in Table 5.2c, a higher proportion of male applicants report various programming languages, concepts, and applications. In addition to the terms seen in Table 5.2c, male applicants mention “app”, “code”, “oop”, “arduino”, “ture”, “c#”, “php”, and “javascript”, more often than female applicants.

Table 5.2c suggests that female applicants were more likely to learn how to program through courses and workshops. This is indicated by words including “mark”, “help”, "attend", "tuition", "workshop", specific high school CS courses, and names of after-school coding workshops (some of these words can be seen in the table on the right in Table 5.2c). Manual inspection revealed that "mark" referred to earning a mark in a course and "attend" referred to attending a programming workshop or event. On the other hand, male applicants mentioned "project" and "selflearn" more often.

Similar gender differences in programming experiences were found in applicants to individual programs, including Computer and Software engineering.

Extracurricular activities: Extracurricular activities of male applicants tend to display a technical focus. As can be seen in Table 5.2d, male applicants mention words including, "robot", "comput", "develop", "websit", "video", and names of specific math and robot competitions, more often than female applicants. In addition to technical extracurricular activities, male applicants mention activities related to sports and music (indicated by words including "coach", "intramur", and "band" in Table 5.2d).

On the other hand, female applicants list a wide breadth of experiences. They mention extracurricular activities related to art, leadership, dance, music, community welfare, event planning, drama, and teaching, more often than male applicants (words related to some of these activities can be seen in Table 5.2d).

Jobs: When describing jobs students held through high school, male applicants were more likely to mention terms that imply technical work (indicated by words including, "comput", "repair", "mainten", and "technician" in Table 5.2e), manual labour, or sports. On the other hand, female applicants were more likely to mention terms that imply jobs including, customer service and teaching (see Table 5.2e).

Reading Interests: More male applicants report reading technical content such as research papers and articles, and more female applicants report reading novels and material with a societal focus. In addition to the words seen in Table 5.2f, male applicants mention words including, "product", "concept", "mechan", "softwar", "machine", "idea", and "develop", and female applicants mention words including, "courag", "encourag", "child", "parent", "world", "death", "relationship", and "suffer".

Additional Information: We see a difference in word choice between male and female applicants when answering a question with no restrictions on the content of their answer. As can be seen by the tokens listed in Table 5.2g, male applicants use this response to highlight their technical interests. In addition to the technical tokens seen in Table 5.2g, male applicants mention tokens such as "math", "java", and "machin" more often than female applicants. Besides technical interests, male applicants are more likely to mention tokens related to sports.

On the other hand, female applicants use this response to highlight their wide breadth of experiences ranging from leadership to artistic pursuits. They also use the response to (a) emphasize their desire to contribute to society (indicated by words including, “volunt”, “peopl”, “communiti”, “educate”, and “chang”, some of which can be seen in Table 5.2g), (b) highlight the influence of family and other role models in their decision to study engineering (shown by words such as “famili”, “parent”, “friend”, and “encourag”), and (c) re-iterate their “passion” and “love” for engineering. Besides, female applicants are more likely than male applicants to list the skills (for example, “hardwork”, “communic”, “creativ”, “activ”, and “timemanag”) that make them suitable for the engineering profession.

5.1.4 Discussion

Since most of the past research on the gender gap in engineering major choice is based on small datasets (collected through surveys and interviews), the goal of this section is to provide data-driven insights into the differences between male and female applicants and their motivations to apply to engineering. Below, we discuss our observations regarding the similarities and differences between male and female applicants.

Observation #1: Regardless of gender and program of choice, the most commonly mentioned words in response to *Why are you interested in joining engineering or the university?* are terms from science and technology (Tables 5.2a and 5.2b). Both male and female applicants mention that they want to study engineering because of technical interest. Past research agreed with our observation. Surveys of both men and women found technical interest and aptitude to be the most important and commonly stated reason behind STEM major selection [16, 143, 342, 313, 339, 197].

However, we found a gender difference in the expression of technical interest. While male applicants reported technical interest by mentioning practical applications of STEM including “robot”, “machines”, and “games”, female applicants mentioned their “love” for “science”, “innovation”, and “research”. Applicants’ programming experience and reasons to apply to the university agreed with this gender difference in application orientation (Tables 5.2c and 5.2b). For example, male applicants mentioned multiple programming languages, applications, and projects, whereas female applicants mentioned attending courses and workshops and basic programming concepts and web development. In addition, male applicants mentioned the university’s startup culture as a reason to apply, more often than female applicants.

Past studies agree with this observation and attribute multiple reasons for this difference. For example, some studies found men to be interested in the practical and investigative components of study, and women to show a stronger creative and social bent [308]. While women were more interested in in-class activities [128], men were more interested in extracurricular activities and competitions [46, 169]. Moreover, more men participated in science extracurricular activities and held a better attitude towards it [33, 169]. Women, on the other hand, were less confident with complex computer tasks [39] and showed less interest in technology-related extracurricular activities [304, 286]. These differences have been attributed to gender differences in STEM exposure and experience (during childhood and adolescence) as well as peer and parental support and encouragement [85, 39]. Researchers speculate that these differences are further reinforced by negative stereotypes against women's STEM abilities and the under-representation of women in STEM extracurricular activities [339].

Another explanation for the gender difference in the mention of practical applications could be a gender difference in the motivation for these tasks. A study found very few girls to be interested in writing code to create a robot, but many more to be interested in writing code to create music, art, or a medical device [87]. Similarly, in a study with 437 students, girls expressed less interest in extracurricular activities consisting of science experiments, and more interest in activities related to the environment or people [169]. Another research studying 268 science projects made by students from Grade 1 to 6 found that, irrespective of grade level, boys tended to choose projects in physical sciences, and girls in biological and social sciences [1]. Overall, women were more interested in STEM applications that they considered *worth pursuing*, for example, activities with communal orientation. Therefore, in order to attract more girls towards engineering activities, and in turn, majors, school STEM curriculum could leverage girls' existing interests and conduct STEM activities with more creative and communal goals [340, 337]. While the reasons stated above provide some explanations for the observation, the difference warrants further research, maybe in the form of interviews.

Observation #2: The overarching gender difference throughout the analysis is that male applicants differentiate themselves through depth of technical experience, whereas female applicants through a breadth of experiences in various fields. To study engineering, all applicants must demonstrate knowledge in mathematics and sciences through their academic work. However, we see male applicants differentiating themselves by highlighting their initiative to acquire more technical skills through their work experience, extracurricular activities, and reading interests. When asked why they want to join engineering or the university, male applicants highlight their technical interests and the university's technical reputation. Moreover, even their response to the question with no restrictions has a technical focus.

On the other hand, female applicants differentiate themselves by demonstrating a wide range of experiences and capabilities. This is suggested by their responses to all the questions on the application form (Table 5.2). Female applicants take on various kinds of jobs including service, teaching, and research, their extracurricular activities place an emphasis on leadership and artistic pursuits, and they choose to discuss more non-technical reading material. Moreover, in their response to the question with no restrictions, they mention their passion for engineering, but also emphasize other areas of interest.

Past studies support our findings. Psychologists suggest that females have multiple cognitive strengths and are highly skilled in both verbal and math domains, in comparison to males who demonstrate higher math relative to verbal ability [340, 339, 329, 181]. These multiple cognitive strengths might have led female students to pursue, and thus, report a variety of interests. Besides, pursuing a variety of interests might have reduced the time allocated to technical endeavours, in turn forcing female applicants to highlight breadth instead of depth; studies note that women entered engineering with greater anxiety and less STEM preparation than men [113, 137]. On the other hand, a study analyzing self-reported performance on arithmetic tasks found that men tend to boast about their performance, whereas women generally under-report it [269]. A combination of these reasons might have led to our observation.

Observation #3: Unlike male applicants, female applicants mention motivations other than technical interests to join engineering. The first of two is family influence. Female applicants are more likely than male applicants to mention personal and family influences in their decision to study engineering (Tables 5.2a and 5.2g). Our observation aligns with past studies, which found encouragement and guidance from parents and teachers to be one of the most important factors why students, especially women, chose STEM careers [290, 312, 143, 230, 139, 16, 85, 220, 206, 85, 339]. Women who had access to role models were not only more likely to choose STEM careers, but also had higher success beliefs and an increased sense of compatibility with STEM careers [57, 287, 353]. However, parental pressure could have also resulted in this difference; women were found to be twice as likely as men to agree to their parent's wishes to pursue STEM education [158].

Secondly, female applicants show a stronger desire to contribute to society. In addition to family influence, they mention improving the world around them as a reason to join engineering, more often than male applicants (Tables 5.2a and 5.2b). This difference is seen in applicants to all programs, including applicants to programs such as Biomedical and Environment that focus on helping others, as well as applicants to Mechanical and Electrical that are farthest removed from directly working for people. Female applicants choose to reiterate their altruistic tendencies in their responses to the question with no restrictions (Table 5.2g). In addition, female applicants' altruistic tendencies are evident in their past work experiences, where they work as a "teacher" or "counsellor" more often than male applicants, their extracurricular activities, where they participate in community-related activities such as volunteering and fundraising more often, as well as their reading interests, which have a greater societal focus.

Past studies confirm that women have more altruistic tendencies than men [308, 87, 99, 1, 197, 169, 290]. When asked about careers, men value money and fame, and women prefer occupations that allow them to work with people and make positive contributions to society [169, 308]. Altruism explains why women prefer socially oriented professions, including medical science and environment [124, 88, 308, 1, 68, 278, 68]. However, it is also a commonly stated reason for them not joining engineering programs [87, 99, 1, 340, 337, 308, 197, 169, 290]. Abundant stereotypes lead many women to believe that math-intensive careers are inconsistent with their desire to work for and with people [88, 87, 99]. A study found that women who aspire towards math-related or engineering careers place a lower value on society-oriented job characteristics [99]. Our observation indicates otherwise and suggests that in order to attract more women to study engineering, it must be presented as a profession that can help others. Programs that increase students' math task values and raise awareness about the applicability of STEM (especially in communal goals) have been shown to increase students', especially women's, interest in STEM careers [338, 13, 221, 125, 207].

In addition to the aforementioned reasons, a higher proportion of female applicants refer to joining engineering as a "challenge" or an "opportunity" (Tables 5.2a and 5.2b). Moreover, female applicants are more likely to highlight their "love" of science, along with the qualities that make them suitable for the profession. Since engineering programs are male-dominated and have a pro-male ability bias, female applicants may join them as a challenge and feel compelled to provide justifications about their suitability [290, 339, 197]. On the other hand, male applicants refer to joining engineering as their "goal" or "childhood dream".

5.1.5 Summary and Conclusions

In this section, we presented a large-scale text mining analysis of the gender differences in engineering applicants. Our analysis was enabled by a unique dataset of over 30,000 undergraduate applications to the engineering faculty of a large North American university. We used syntactic text analysis techniques to infer why male and female applicants were interested in studying engineering, and how they differed in their reading interests, extracurricular activities and programming experience. To the best of our knowledge, this is the first large-scale work that measures gender differences in motivations and interests of engineering applicants. The study provides data-driven evidence for past findings as well as provides new insights into gender differences in students who want to join engineering programs.

Our analysis revealed that both male and female applicants wanted to join engineering due to technical interests. While both mentioned terms from science and technology, male applicants were more likely to mention practical applications of STEM, including robots and machines. Moreover, male applicants differentiated themselves through technical depth and female applicants differentiated themselves through a breadth of experience in various fields, including arts and volunteering. In addition, female applicants displayed a greater desire to serve society and were more likely to mention interpersonal relationships and familial role models when discussing their engineering goals.

We infer that to attract more female students to study engineering, it must be presented as a profession that can help others and allow for a broad range of careers and learning opportunities. We believe that the message to female students should not just be that they can do it, but that they should want to do it because engineering is an excellent fit for their values and priorities. Emphasis on real-life STEM applications during high school and first-year engineering courses that show the wide-ranging impact of technology on society may help change this messaging. In addition, our results suggest that a key part in fostering this new image of engineering lies in encouragement from family and role models who practice engineering. Moving on to directions for future work, similar data-driven analyses of applications to graduate school and non-STEM programs may provide insights into their gender imbalance. Moreover, correlating depth and breadth of expression at the time of admission to academic and career success may be another interesting direction for future work.

5.2 Gender differences in applicants' perceptions of co-operative education

5.2.1 Motivation

Section 5.1 studied gender differences in students' motivations to join engineering programs. Since all engineering programs at the university have mandatory co-op, this section focuses on applicants' reasons to join co-op programs in particular, and measure any gender differences in students' perceptions of co-op. In the past, the benefits of co-op education have been studied by surveying students already enrolled in co-op programs [5, 265]. In contrast, this analysis focuses on understanding what prospective students think about co-op. Adjusting recruiting material and outreach programs based on the results of this section may highlight the aspects of co-op that potential female students may find desirable, and in turn attract more female students and increase diversity. To the best of our knowledge, there is no previous work on investigating gender differences in prospective students' expectations from co-op.

5.2.2 Data and Methods

This analysis is enabled by the Admissions dataset (described in Section 3.1), and in particular, the 33,763 non-blank applicant responses to the question "Tell us about your reasons for applying to this university". We apply the text mining method described in Sections 4.2.2 and 4.2.3 to this field, where we first separate applicants' reasons involving co-op from other reasons for joining the university, and then compare the words used by male and female applicants while talking about co-op.

The parser described in Section 4.2.2 first applies some pre-processing to each applicant response. In this process, among other operations, it converts alternative forms of the word co-op to "coop". The parser then filters and retains the sentences containing the token "coop", and finally, converts the extracted sentences to tokens. Figure 4.2 summarizes these steps and shows the set of tokens extracted from an applicant's response, possibly related to the applicant's reasons for joining a co-op program. Next, we compare the frequency of these tokens in male and female applicant responses using the term frequency analysis described in Section 4.2.3. We present the frequent terms present in these sentences and the gender differences in their frequency. Common English words with significant differences are excluded from the report for brevity.

The method described above has the following limitations. First, it was assumed that any sentence containing the term "coop" (or one of its alternative forms) reflects an applicant's opinion on co-op. This may lead to some false positives (sentences that mention "coop" but focus on a different topic) and false negatives (sentences that discuss co-op without mentioning the word). Manual inspection of a random sample of 50 responses revealed one false positive and four false negative sentences. False positives occurred when applicants mentioned their program of choice, such as "I am applying to the Biomedical Engineering co-op program because it is unique". False negatives occurred when students wrote multiple sentences about co-op but did not include the term "coop" in every sentence. In the random sample that was manually inspected, these additional sentences paraphrased topics mentioned by the sentences containing the term "coop", meaning that removing these sentences resulted in minimal information loss.

Next, it was also assumed that any sentence containing the term "coop" (or one of its alternative forms) specifies a reason why an applicant is interested in a co-op program, instead of why they are *not* interested in it. To confirm this, each sentence containing the term "coop" (or one of its alternative forms) was inspected by a sentiment analyzer[3]. Examples of synthetic sentences tagged as positive include: *"Participation in Co-op will help me find a full-time job after graduation"* and *"I am applying to the university's Chemical co-op program"*. Examples of synthetic sentences tagged as negative include: *"Learning the theory feels purposeless without the chance to apply it, and co-op would solve this"* and *"I want to join the university in spite of its co-op program"*. The analyzer assigned a positive score to 99% of the sentences in the dataset. Manual inspection of all the sentences with a negative score revealed that they contained a strong negative word (for example, "purposeless" in the sentence above), but this word was not directed towards co-op.

The same problems were seen in sentences extracted using the Question Answering technique available as an open source API[4]. In fact, the Question Answering API, even with multiple variations of the question "What do you think about co-op?" as input, missed more sentences and documents than the simple pre-processing *Filter* step.

5.2.3 Results

Frequency analysis of applicant responses to the question "Tell us about your reasons for applying to this university" shown in Section 5.1.3 indicates that female applicants mention

[3] https://www.nltk.org/api/nltk.sentiment.html
[4] https://github.com/allenai/bi-att-flow

“coop” 5% more often than male applicants. This difference is statistically significant with a p-value less than 0.001.

To understand why more female than male applicants mention co-op as a reason to join the university, we focus on sentences that contain the token “coop” (i.e., we apply the *Filter* method mentioned in Section 4.2.2). Below, we categorize and list the frequent tokens in these sentences. This list does not include the frequently occurring English words and words common in the context of university applications (e.g., “appli”, “educ”, “provid”, “student”). Applicants' mention co-op in the following contexts:

1. *To gain knowledge and skills:* “learn” (14 percent of applicants who mention “coop” in a sentence mention “learn” in the same sentence), “interest” (10%), “knowledge” (9%), “skill” (7%), “valuable” (6%)

2. *To gain work experience:* “work experience” (14%), “work” (14%), “field” (9%), “many jobs” (5%), “research” (4%)

3. *To gain practical experience:* “practical experience” (8%), “practic” (6%), “real-world” (4%). Upon manual inspection, it was found that applicants described co-op as an opportunity to apply the “theory” (2%) learned in the “classroom” (3%) to solve real-world problems.

4. *Reputation and size of the co-op program:* “reput” (13%), “best” (9%), “world” (7%), name of country where the institution is located (7%), “rank” (5%), “renown” (4%), “largest” (4%)

5. *Career prospects:* “career” (9%), “future” (8%), “graduat” (7%), “job” (6%), “employ” (4%), “degree” (4%), “placement” (3%).

Other tokens mentioned alongside “coop” include: “unique” (5%), “connect” (3%), “compani” (3%), “explore” (2%), “network” (2%), “finance” (1%), “tuition” (1%), and “entrepreneur” (1%).

To understand whether male and female applicants perceive co-op differently, we list terms with statistically significant differences in frequencies between the two groups. Initially, we analyzed gender differences in applicants to each program separately, but later we observed the trends in each program to be similar to those displayed by all applicants. Therefore, we omit the per-program details for brevity.

Table 5.3: Differences in frequencies between tokens mentioned by male and female applicants (in responses containing sentences with "coop")

Token	Male	Female	Δ
reput	14%	11%	3%***
reason	10%	7%	3%***
best	10%	8%	2%***
excel	8%	6%	2%***
(city)	9%	7%	2%**
industri	4%	2%	2%***
compani	4%	2%	2%***
world	11%	9%	2%**
(country)	7%	6%	1%**
employ	4%	3%	1%**

Token	Female	Male	Δ
opportun	26%	21%	5%***
learn	16%	14%	2%**
career	10%	8%	2%**
attract	6%	4%	2%***
explor	3%	1%	2%***
practicalexperi	9%	8%	1%**
uniqu	6%	5%	1%***
think	3%	2%	1%***
love	3%	2%	1%***
different	4%	3%	1%***

Table 5.3 shows the differences in frequencies of tokens mentioned by male and female applicants. Table 5.3 shows the top 10 tokens that are mentioned significantly more frequently by male applicants than by female applicants (on the left), and vice versa (on the right). The lists are sorted by the difference in frequencies, abbreviated Δ, and the asterisks indicate the strength of the statistical significance of the difference. For example, the first row in the table on the left in Table 5.3 shows that, when talking about "coop", male applicants mention "reput" 3% more often than female applicants and the p-value of this difference is less than 0.001.

As can be seen in Table 5.3, male applicants, more often than female applicants, mention tokens related to reputation, size, and the companies that participate in the co-op program (tokens suggesting this include "reput", "best", "renown", "prestig" "largest", "world", name of the country where the institution is located, names of employers that participate in the co-op program, etc.).

On the other hand, female applicants mention tokens related to gaining knowledge ("learn", "knowledge") and practical experience ("practical experience", "practice", "workplace", "theory", "classroom", "field") and exploring a variety of career options ("explore", "different", "various", "variety", "divers", "options", "paths") including research, more often than male applicants. Some of these tokens can be seen in Table 5.3.

5.2.4 Discussion

Analyzing applicants' reasons to apply to the university, specifically its engineering programs with mandatory co-op, led to two main observations. The first revealed a difference in importance that male and female students gave to a co-op program and the second indicated a gender difference in students' perception of co-op. While further research is required to understand the reasons behind the differences found, our study is the first to identify gender differences in students' views of co-op programs.

Observation #1: Female applicants mentioned co-op 5% more often than male applicants as a reason for applying to the university. Thus, institutions wishing to increase female enrollment in engineering may benefit from emphasizing co-op in their outreach efforts.

Observation #2: Male applicants were more likely to mention the size and reputation of the institution's co-op program, whereas female applicants were more likely to talk about co-op as an opportunity to gain knowledge, practical experience, and try a variety of career options. Outreach programs and recruitment material aimed at attracting more female students should therefore emphasize these benefits.

In addition to informing recruitment material, knowing what prospective male and female students think about co-op can help institutions identify inconsistencies between expectation and reality and manage students' expectations. For example, female applicants stated that they wanted to participate in co-op to learn new technical skills and apply theories learned in the classroom to real-world problems. However, studies found that a majority of co-op positions available during junior years were not always directly related to students' field of study [61]. Our findings can help institutions identify such inconsistencies between expectation and reality and help students manage their expectations, in turn increasing their satisfaction with co-op and retention in the program.

Being aware of incoming students' mindsets towards co-op can increase the likelihood of meeting their expectations. Institutions may wish to align the co-op program to suit the needs of both male and female students. For example, female applicants stated that they want to explore different jobs in their field in order to make an informed career choice. To accommodate this, institutions may consider recruiting employers who are willing to rotate students among different teams or business units during a work term. On the other hand, male students stated that they want to work in large technology companies during their co-op work terms. Thus, institutions may want to organize information sessions where representatives from these companies explain their talent needs. In addition, this may also serve as a platform for aspiring co-op students to meet senior students who have worked for these companies in the past and benefit from their recruitment strategies. Meeting students' expectations and incorporating their needs into co-op programs may not only retain current students but also attract diverse talent.

Moving on, since the observation reveals what female applicants expect from their co-op experience, co-op employers wishing to diversify their talent pool can adjust their job descriptions to attract more female applicants. For example, our results suggest that female students perceive co-op as a means to gain technical, practical, and workplace skills. Therefore, an additional section in co-op job descriptions, which generally only contains sections on required skills and job responsibilities [195, 59], describing *skills that students can expect to learn on the job* may encourage female students to apply to these positions.

A combination of reasons may explain the *two* observations stated above. For example, past studies found that many women prefer a kinesthetic learning style over visual, auditory, or read and write styles [344]. This may explain why female applicants emphasize co-op and practical experience more than male applicants. Other studies suggest that employers have a higher hiring standard for women [268]. Knowledge of this gender bias may explain why female applicants emphasize gaining work experience [107, 182, 232, 269].

The results of Section 5.1 may also provide some explanations for these observations. For example, in Section 5.1, we found that, on average, female applicants to engineering have less technical experience in terms of part-time jobs and extracurricular activities. This difference in technical experience may explain why female applicants mention co-op more often than male applicants, with female applicants viewing co-op as a way to gain these skills and work experiences. Similarly, Section 5.1 found female applicants to possess and report a wider variety of interests and experiences than male applicants. This may explain why female applicants mention a desire to try different career options through co-op.

5.2.5 Summary and Conclusions

In this section, we presented gender differences in engineering applicants' perceptions of co-op. As part of the application process to the engineering faculty of a large North American university with mandatory co-op, more than 33,000 applicants answered the question: "Tell us about your reasons for applying to this university". We applied text mining techniques on applicants' responses to this question to extract and identify gender differences in their reasons to apply to a co-op program. Our study is the first to examine prospective students' perceptions of co-op and how they may differ by gender.

Our analysis suggested that female applicants mentioned co-op as a reason to join the university more often than male applicants. In addition, while female applicants wanted to join the co-op program to learn new skills, try a variety of career options, and gain practical work experience, male applicants wanted to join it to leverage the reputation and size of the institution's co-op program.

These findings are relevant to students, institutions, and co-op employers. Institutions may use prospective students' perceptions of co-op to inform outreach and recruitment efforts and attract diverse talent. Additionally, institutions and co-op practitioners may use these findings to identify gaps between perceptions and reality and either manage students' expectations early in the co-op process or adapt the process to meet students' needs, in both cases, increasing satisfaction and reducing attrition. Finally, our findings are beneficial for co-op employers as it may help them attract diverse talent. Since this study focuses on gender differences in engineering applicants' perceptions of co-op programs, investigating the perceptions of co-op of other underrepresented groups (for e.g., racial minorities) and applicants to other faculties and universities may provide additional insight into promoting co-op programs.

5.3 How to increase the number of female engineering applicants from high schools

5.3.1 Motivation

Past research, along with Sections 5.1 and 5.2, studied gender differences in students' personal considerations towards joining engineering programs. Studies found gender differences in students' interest in the field, academic performance in high school STEM courses, sources of influence (including parents and teachers), and career perceptions, to greatly affect the proportion of women interested in engineering majors [339, 85, 290]. For instance, Section 5.1 along with other studies suggested that a mismatch between female applicants' altruistic values and the engineering profession's image may be contributing towards the 77% male-dominated applicant pool of engineering [339, 169, 124, 88, 308, 1, 278, 68]

Since much of the past research on STEM recruitment focuses on students' individual considerations, researchers note that less work has been done to understand the effect of a student's surrounding system (especially their high school contexts) on their post-secondary choice of major [136, 191, 256, 105]. For example, Lee [191] suggested that more work is needed to understand situations "in which participants are acted upon by a surrounding system and have little agency to change their course". Thus, in this section, we conduct a novel data-driven study to analyze the high school contexts of engineering applicants and identify the aspects that affect female students' interest in engineering. We do this by identifying, at an aggregate level, the unique characteristics of schools (and their applicants) that produce many female engineering applicants. Identifying the characteristics of high schools that promote its female students to apply to engineering programs may help us provide actionable insights into the effect of students' surrounding systems and decrease the gender gap in the engineering applicant pool.

5.3.2 Data and Methods

Our analysis is enabled by two unique datasets: (1) the Admissions dataset, which contains data extracts of more than 33,000 applications to the undergraduate engineering programs at the university (between 2013 and 2016). For each application, the dataset includes the student's responses to questions about their interests and background (described in Section 3.1) and the name of their high school, and (2) the High school dataset, which contains aggregate demographics and academic performance statistics for all public high schools in the province of Ontario, Canada, where this university is located (details in

Section 3.2). To answer the research question, *Why do some high schools produce more female engineering applicants than other high schools?* (mentioned in Figure 1.1), we cross-reference these two datasets and obtain a list of 670 public high schools across Ontario that send 17,814 applications to the university between 2013 and 2016. As explained in Section 3.2, in our analysis we only consider schools with greater than 25 applicants per year or greater than 100 applicants over the four years.

We start with statistical analysis, where we use the Pearson correlation coefficient to identify correlations between the proportion of female engineering applicants from a given high school and the school's demographics and academic performance metrics. Along with the correlation coefficient, we report the result of a statistical significance test, which indicates whether the correlation coefficient found is statistically significant. Details about this method can be found in Section 4.1. For this analysis, we used schools that sent at least 25 applications in 2016 since this was the only overlapping year between the admissions and the high school datasets (students who applied to engineering in 2016 wrote the Grade 9 OSSLT exam in 2013 and the Grade 10 Math exam in 2014).

Next, to distinguish between schools with many versus few female engineering applicants based on what the students from these schools said in their applications, we use the text mining method explained in Section 4.2. For this analysis, we consider the 31 schools that produce at least 100 engineering applicants between 2013 and 2016. Using the method described in Section 4.2.1, we first identify the schools that produce a much higher and lower proportion of female engineering applicants and label them M_schools and F_schools respectively. Each group contained eight schools and more than 1000 applicants each. Then, for every non-blank student response to each of the seven questions in the Admissions dataset, we use the parser described in Section 4.2.2 to convert the response into a set of tokens. Finally, for each of the seven text fields in the Admissions dataset, we apply the term frequency analysis described in Section 4.2.3. To understand the differences between M_schools and F_schools, we compare the responses of male applicants from M_schools to those of male applicants from F_schools. Separately, we compare the responses of female applicants from M_schools and to those from F_schools. We report the significant differences in word frequencies for all questions expect reading interests because they show no difference between M_schools and F_schools. In addition, we exclude common English words with significant differences for brevity.

5.3.3 Results

In this section, we present the differences between schools that produce many applications from female students and those that produce few applications. Particularly, we present the results of applying statistical and text analysis methods to identify (a) correlations between a school's demographics and academic performance metrics with the percentage of female engineering applicants it produces, and (b) unique characteristics of schools that produce many female engineering applicants based on what students from these schools said in their applications.

5.3.3.1 Statistical Analysis

Table 5.4 shows the correlation coefficients (and the p-values) of the relationships between the percentage of female engineering applicants and the high school characteristics. The only variable with a statistically significant correlation coefficient in Table 5.4 is the gender gap in Grade 10 Math scores. This suggests that schools where female students perform better than male students in Math provincial exams produce a higher proportion of female engineering applicants. The magnitude of the correlation coefficient (0.32) indicates a medium positive association.

Table 5.4: Correlation of % Female Applicants from a school with other metrics of the schools

Category	Metric	Correlation Coefficient	P-Value
Demographics	% of students who live in lower-income households	0.16	0.28
	% of students whose parents have some university education	0.10	0.52
	% of students whose first language is not English	0.16	0.29
	% of students who are new to Canada	0.03	0.83
Performance	% of Grade 10 students who passed the English exam	0.11	0.48
	% Female - % Male students who pass the English exam	0.15	0.30
	Average score of Grade 9 students in the Academic Math exam	0.14	0.34
	Average score of Grade 9 students in the Applied Math exam	0.22	0.14
	Female - Male Average Math Scores	0.32	**0.03**
Others	% of engineering applicants	0.01	0.94
	Distance from university	-0.21	0.15

5.3.3.2 Text Analysis

Table 5.5 shows a few tokens with frequency differences between the responses of male applicants from M_schools and F_schools for six questions [5]. For example, Table 5.5a shows tokens that are mentioned significantly more frequently by male applicants from M_schools than by male applicants from F_schools (on the left) and vice versa (on the right), when discussing their engineering interests and goals. The lists are sorted by the difference in frequencies (abbreviated Δ) and the asterisks indicate the statistical significance of the difference. Table 5.6 follows the same format as Table 5.5. It shows word frequency differences in the responses of female applicants from M_schools and F_schools.

Below, we discuss the word frequency differences in the six text responses of applicants from M_schools and F_schools. However, before that, it is important to note that we found similar gender differences as found in Section 5.1 in applicants from M_schools and in applicants from F_schools. We compared the responses of male and female applicants from M_schools (and F_schools) and found words similar to those seen in Table 5.2. Therefore, it is important to keep in mind that the differences between M_schools and F_schools discussed below occur in addition to the gender differences listed in Section 5.1.

Engineering Interests and Goals: Both male and female applicants from M_schools and F_schools mention technical interests. Table 5.5a shows that male applicants from M_schools mention "robot" 7% more often than male applicants from F_schools. On the other hand, male applicants from F_schools mention specific robotic challenges and programming languages (for example, "python") more frequently than male applicants from M_schools. Similarly, Table 5.6a shows that female applicants from M_schools mention technical extracurricular activities, including specific robotics and science competitions, more often than female applicants from F_schools. However, female applicants from F_schools mention computer programming more often than female applicants from M_schools. Overall, all groups mention interests related to robotics, computing, and science.

Notably, Tables 5.5a and 5.6a indicate that applicants from M_schools are more likely to mention collaborative and competitive STEM activities. Male applicants from M_schools mention "team", "mentor", and "compet" more often than male applicants from F_schools. Furthermore, female applicants from M_schools mention specific competitions, "network", and "group" more often than female applicants from F_schools (some of these tokens can be seen in Tables 5.5a and 5.6a).

[5] We observed no difference in the reading interests of applicants from M_schools and F_schools. Thus, their results have not been included in Table 5.5 or Table 5.6

Table 5.5: Differences in frequencies between tokens mentioned by *male* applicants from M_schools and F_schools

(a) Engineering Interest and Goals

Token	M_schools	F_schools	Δ
team	21%	12%	9%***
passion	25%	18%	7%***
robot	25%	18%	7%***
changetheworld	26%	20%	6%***
goal	28%	22%	6%**
mentor	3%	1%	2%***

Token	F_schools	M_schools	Δ
parent	8%	5%	3%**
scientificpotential	5%	2%	3%***
fun	4%	2%	2%***
(robotic challenge)	2%	0%	2%***
python	2%	0%	2%**

(b) Reasons to apply to the university

Token	M_schools	F_schools	Δ
coop	69%	61%	8%***
experi	35%	27%	8%***
degre	28%	22%	6%**
job	10%	6%	4%***
network	5%	2%	3%***

Token	F_schools	M_schools	Δ
math	33%	24%	9%***
physic	22%	13%	9%***
thrill	1%	0%	1%**
pride	1%	0%	1%**
(university's startup incubator)	1%	0%	1%***

(c) Programming Experience

Token	M_schools	F_schools	Δ
visualbas	17%	4%	13%***
java	73%	61%	12%***
c	16%	5%	11%***
(high school CS course)	11%	0%	11%***
ap	10%	1%	9%***
team	10%	3%	7%***
(IB high school course)	4%	0%	4%**
selflearn	2%	0%	2%*

Token	F_schools	M_schools	Δ
python	52%	18%	34%***
c++	32%	15%	17%***
(high school CS course)	6%	0%	6%***
(CS course project)	3%	0%	3%*
pygam	3%	0%	3%**
text	2%	0%	2%*
exam	2%	0%	2%*

Table 5.5: Differences in frequencies between tokens mentioned by *male* applicants from M_schools and F_schools, continued

(d) Extracurricular activities

Token	M_schools	F_schools	Δ
robot	21%	11%	10%***
team	44%	35%	9%***
school	55%	47%	8%***
lead	18%	12%	6%***
design	13%	8%	5%***
communic	9%	5%	4%**

Token	F_schools	M_schools	Δ
orchestra	18%	7%	11%***
(community service club)	8%	0%	8%***
(youth business organization)	14%	9%	5%***
(university)	7%	3%	4%***
debat	6%	2%	4%***
fundrais	6%	3%	3%***
footbal	5%	2%	3%***

(e) Jobs

Token	M_schools	F_schools	Δ
assist	17%	11%	6%**
comput	3%	1%	2%**
hockey	2%	0%	2%***
associ	2%	0%	2%***
server	2%	0%	2%**

Token	F_schools	M_schools	Δ
regist	4%	1%	3%**
babysitt	1%	0%	1%**
bus	1%	0%	1%*
coffe	1%	0%	1%**
outdoor	1%	0%	1%*

(f) Additional Information

Token	M_schools	F_schools	Δ
(STEM program)	7%	0%	7%***
collabor	3%	0%	3%***
librari	3%	1%	2%**
mentor	3%	1%	2%**
(robotic competition)	2%	0%	2%**
changetheworld	2%	0%	2%**

Token	F_schools	M_schools	Δ
IB	7%	2%	5%***
player	3%	0%	3%***
python	2%	0%	2%**
(math competition)	3%	0%	3%**
(after-school math program)	1%	0%	1%*
orchestra	1%	0%	1%*
reasoningskil	1%	0%	1%*
privat	1%	0%	1%*

Table 5.6: Differences in frequencies between tokens mentioned by *female* applicants from M_schools and F_schools

(a) Engineering Interest and Goals

Token	M_schools	F_schools	Δ
pursu	34%	22%	12%***
extracurricular	4%	0%	4%***
network	4%	0%	4%***
(robotic competition)	2%	0%	2%**
(science competition)	2%	0%	2%**

Token	F_schools	M_schools	Δ
friend	11%	4%	7%***
spoken	3%	0%	3%*
elementaryschool	3%	0%	3%*
build	2%	0%	2%*
algorithm	2%	0%	2%*
IB	2%	0%	2%*

(b) Reasons to apply to the university

Token	M_schools	F_schools	Δ
bachelor	10%	1%	9%***
changetheworld	7%	1%	6%***
disciplin	7%	2%	5%***
financialsupport	2%	0%	2%*
seniorposition	2%	0%	2%*

Token	F_schools	M_schools	Δ
canada	16%	7%	9%***
strongfoundat	8%	3%	5%**
rank	5%	1%	4%**
cultur	4%	0%	4%*
IB	2%	0%	2%*
startup	2%	0%	2%*

(c) Programming Experience

Token	M_schools	F_schools	Δ
visualbas	13%	0%	13%**
2d	9%	0%	9%*
team	3%	0%	3%**

Token	F_schools	M_schools	Δ
python	44%	9%	35%***
algorithm	13%	0%	13%*
mark	8%	0%	8%*

Table 5.6: Differences in frequencies between tokens mentioned by *female* applicants from M_schools and F_schools, continued

(d) Extracurricular activities

Token	M_schools	F_schools	Δ
lead	45%	31%	14%***
robot	15%	6%	9%***
peer	7%	0%	7%***
mentor	9%	3%	6%***
ontario	9%	3%	6%**

Token	F_schools	M_schools	Δ
(community service club)	9%	1%	8%***
debat	9%	3%	6%**
(university)	7%	2%	5%**
travel	4%	0%	4%**
(STEM summer program)	5%	1%	4%**
hobbi	5%	1%	4%**

(e) Jobs

Token	M_schools	F_schools	Δ
summer	11%	3%	8%***
computerservic	3%	0%	3%**

Token	F_schools	M_schools	Δ
aid	5%	0%	5%*

(f) Additional Information

Token	M_schools	F_schools	Δ
mechan	8%	1%	7%***
robot	7%	1%	6%***
(STEM program)	5%	0%	5%***
growth	4%	0%	4%**
team	4%	0%	4%**
joy	4%	0%	4%**
financ	4%	0%	4%**

Token	F_schools	M_schools	Δ
chanc	9%	2%	7%**
hobbi	6%	0%	6%*
capabl	6%	0%	6%**
advic	4%	0%	4%*
math	4%	0%	4%*

Moving on, applicants from F_schools are more likely to mention personal influence and guidance from family members and friends. Male applicants from F_schools mention “parent”, “father”, and “mother” more often than male applicants from M_schools (some of these tokens can be seen in Table 5.5a). Female applicants from F_schools mention tokens including, “friend”, “spoken”, “brother”, “advice”, and “recommend” more often than female applicants from M_schools (some of these tokens can be seen in Table 5.6a).

Lastly, applicants from M_schools and F_schools report different motivations to study engineering. Applicants from M_schools refer to pursuing engineering as their “passion” or “goal” and as a means to help the world (see Tables 5.5a and 5.6a). On the other hand, applicants from F_schools want to study engineering because they did well in STEM courses and enjoy STEM activities (indicated by words including “scientificpotential” and “fun” in Table 5.5a and “elementaryschool” and “IB” in Table 5.6a).

Reasons to apply to the university: Applicants from M_schools and F_schools mention different reasons to apply to the university. Applicants from M_schools mention career planning, security, and growth, as reasons to apply to the university more often than applicants from F_schools. To begin with, male applicants from M_schools mention “coop” and “degree” more often than male applicants from F_schools (seen in Table 5.5b). Similarly, female applicants from M_schools mention getting a “bachelor” degree in their “discipline” of choice more often than female applicants from F_schools (Table 5.6b). Additionally, applicants from M_schools wish to join the university to gain work experience (indicated by the word “experi” in Table 5.5b) to make it easier to find a job after graduation (“job” in Table 5.5b and “seniorposition” in Table 5.6b), make connections in the industry (“network” in Table 5.5b), and pay for their education (“financialsupport” in Table 5.6b). In addition, the token “changetheworld” in Table 5.6b (in addition to its presence in the Table 5.5a) suggests that applicants from M_schools have more altruistic motivations than applicants from F_schools.

On the other hand, applicants from F_schools mention their love and aptitude for STEM and the university’s reputation as reasons to apply to the university, more often than applicants from M_schools. This is indicated by words such as “math”, “physics”, “thrill”, and “pride” in Table 5.5b and “canada”, “strongfoundation”, “rank”, “culture”, and “IB” in Table 5.6b. Additionally, both male and female applicants from F_schools mention the university’s entrepreneurship culture as a reason to apply to the university, slightly more frequently than applicants from M_schools.

Programming Experience: Both male and female applicants from M_schools and F_schools mention various programming languages. Applicants from M_schools are more likely to mention “viusalbasic”, “java”, “c”, “php”, “jquery”, “2d”, and “assembly” (some of these can be seen in tables on the left in Tables 5.5c and 5.6c). On the other hand, applicants from F_schools are more likely to mention "python", "c++", "algorithm", "pygam", and "text" (seen in tables on the right in Tables 5.5c and 5.6c). Particularly, applicants from M_schools mention "viusalbasic" 13% more often than applicants from F_schools and applicants from F_schools mention "python" 34% more often than applicants from M_schools. This difference in the kind of programming languages mentioned by applicants from M_schools and F_schools may suggest a difference in the availability of opportunities to learn new and modern languages in the two sets of schools.

Looking into their sources of learning, applicants from both M_schools and F_schools mention high school computer science courses, projects, and exams. However, applicants from M_schools mention other sources of learning, including advanced courses ("ap"), participation in team activities, and self-learning (this can be seen in the tables on the left in Tables 5.5c and 5.6c).

Extracurricular activities: Applicants from M_schools report more technical, collaborative, and competitive activities. Words such as "robot", "team", "lead", "design", "communicate" and "ontario" in Tables 5.5d and 5.6d indicate that both male and female applicants from M_schools report involvement in robotics teams and competitions more often than applicants from F_schools. Moreover, words including "school", "mentor", "team", and "peer" suggest that these students may belong to teams representing their high schools in these competitions.

On the other hand, applicants from F_schools mention a breadth of extracurricular activities and hobbies, including those related to music, community service, debate, summer camps, business, sports, dance, art, skiing, environmental campaigns, public speaking, fashion, multi-cultural events, and travel. Words representing some of these activities can be seen in the tables on the right in Tables 5.5d and 5.6d.

Participation in a wide range of activities, including an expensive STEM summer program and travel (seen in Table 5.6d), may indicate that applicants from F_schools had more exposure than applicants from M_schools. In addition, both male and female applicants from F_schools mention participating in extracurricular programs organized in the university more often than applicants from M_schools.

Jobs: Applicants from both M_schools and F_schools work as office assistants, shop managers, servers, or other aids. Nevertheless, we note a slight difference in tokens that imply technical work, with applicants from M_schools mentioning it more frequently than applicants from F_schools. Examples of such words include "comput" in Table 5.5e and "computerservic" in Table 5.6e. In addition, applicants from M_schools mention "summer" jobs more often than applicants from F_schools.

Additional Information: In this response with no restrictions on content, applicants from both M_schools and F_schools talk about technical interests. As seen in Tables 5.5f and 5.6f, applicants from both M_schools and F_schools mention various STEM subjects, concepts, programs, and competitions. Nevertheless, we note some differences. Male and female applicants from M_schools emphasize collaborative learning by mentioning participation in team competitions (indicated by words including "collabor", "mentor", and "team"). They also mention words such as "robot" and "mechanical" more often than applicants from F_schools. On the other hand, applicants from F_schools emphasize STEM capability and exposure by mentioning "IB" and "private" STEM programs.

Tables 5.5f and 5.6f also indicate that applicants from M_schools differentiate themselves by emphasizing their interest in STEM, while applicants from F_schools talk about both technical interests and hobbies. Applicants from F_schools mention tokens related to a breadth of interests including, STEM, sports, music, history, philosophy, and language. Some words related to these fields can be seen in the tables on the right in Tables 5.5f and 5.6f.

Lastly, as seen before, applicants from M_schools mention altruism and career growth (indicated by tokens including "changetheworld", "financ", and "growth" in Tables 5.5f and 5.6f) and applicants from F_schools mention capabilities and personal influences (indicated by the tokens "capabl" and "advic").

5.3.4 Discussion

The goal of this section is to use data-driven methods to understand why some high schools produce more female engineering applicants than other high schools. Unlike many past works that study personal considerations of students to understand their choice of major, we focus on identifying differences in students' high school contexts. Moreover, we identify these differences by (a) text mining applications for admission to engineering, and (b) analyzing the schools' average math test scores and the gender gap in these scores, metrics that have not been considered by other studies. Below, we discuss our observations and use them to provide data-driven insights into how to increase the number of undergraduate engineering applications from under-represented groups in engineering such as women. However, it is important to keep in mind that these inferences are drawn from the applications received by a single North American university. Additionally, as is the case with all secondary data analyses, further research may be required to establish cause and effect.

Observation #1: We found that more female students apply to engineering from schools where female students outperform male students on Grade 9 provincial Math exams (seen in Table 5.4 in Section 5.3.3.1). Past studies found that intent to pursue a STEM degree is affected by high school math achievement [342, 338, 206, 313]. Moreover, studies found that performance in STEM courses affects students' attitude and future course selection, which in turn affects their choice to enroll and persist in STEM [209, 24]. Thus, performing well in Grade 9 Math may have encouraged female students to take STEM courses in high school and ultimately apply to an engineering program at the university.

Another explanation is related to women's self-efficacy and competence beliefs in STEM. Studies found that students' self-concept of their ability to do well in math and science plays a key role in choosing STEM careers [339, 102, 172, 342, 337]. Students' ability beliefs were found to be affected by their study environment (through their childhood, adolescence, and adulthood) [339, 85, 220, 206]. Past studies found that societal beliefs, stereotypes, and biases related to gender differences in STEM ability reduce women's self-concept in STEM and impact their decisions to pursue STEM careers [102, 339, 313]. A study found both men and women to possess a pro-male STEM bias [110]. In addition, women believed that innate intelligence is needed for success in engineering and that they are less likely to possess these qualities [339]. Girls did not feel capable in math and science [308, 338] and found it more suitable for boys [169]. Past studies have suggested that being stereotyped as less competent by society may increase stereotype threat and create self-efficacy doubts in women, reducing their STEM performance and increasing their tendency to choose non-STEM careers [314, 214].

Observation #2: Applicants from schools with varying proportions of female applicants report no difference in technical interests. In general, applicants report an interest in programming, robotics, and other STEM-related activities. However, both male and female applicants from schools with a higher proportion of female applicants report participating in fewer collaborative STEM activities and competitions. Evidence for this can be found in applicants' descriptions of their engineering interests and goals, programming experience, extracurricular activities, jobs, and other information in Tables 5.5 and 5.6 in Section 5.3.3.2. Related work presents conflicting reports on whether participation in team-oriented STEM activities increases or decreases women's interest in STEM.

A study found that girls who participated in STEM activities showed more interest in STEM, received better grades, and had stronger STEM career aspirations than boys who participated in similar activities [336]. On top of that, some studies indicate that collaboration in STEM activities is particularly helpful for girls. Girls with peer groups who encouraged, endorsed, or exemplified high math and science achievement, had higher math and science motivation [188] and were more likely to see themselves as future scientists [301]. Participating in STEM extracurricular activities with classmates seemed to increase STEM motivation for all students, including women [330]. On the other hand, some studies found that participation in collaborative STEM activities reduced women's interest in pursuing STEM education and careers. A study that surveyed participants of a STEM team competition observed that fewer women than men reported an increased interest in pursuing a STEM career as a result [149]. Various reasons may explain why participation in team-based STEM activities benefits women less than men.

Lack of autonomy and responsibility in STEM team activities can negatively affect girls' attitudes towards STEM [343, 345]. A study that observed interactions within small mixed-gender groups found that boys generally led STEM activities in comparison to girls who usually followed [345]. In another study where interviewers asked students about the benefits of group work, men were more likely to mention explaining the material to others, and women were more likely to mention having the material explained to them [113]. Women were also significantly more likely to feel that other group members undervalued their contributions [113].

Under-representation and lack of social belonging in STEM group activities may be another cause for women to shy away from STEM careers. An evaluation of an international robotics competition for middle and high school students found that boys outnumbered girls by almost 3:1 [149]. A study where college students were randomly assigned to STEM teams found that when women were the minority in a team (less than 25%), they spoke less, were less involved in teamwork, felt less confident, and reported feeling more unsure and worried [84]. Moreover, these students reported lowered engineering career aspirations after the team interaction [84]. This was not the case for women assigned to teams with greater than 75% women. Research suggests that when women are outnumbered by men, they face stereotype threat and perform negatively, thus getting further discouraged from persisting in STEM [302]. Moreover, the lack of social belonging [314], the pro-male STEM stereotypes of team members [169, 301], and everyday sexism in teams [285] may make even interested women reluctant about choosing STEM.

The competitive nature of STEM, which is evident in STEM group activities, can be another reason why women (who participate in these activities) do not want to pursue STEM education [85]. Competition among peers is likely to negatively affect the sense of belonging [158]. A study found that 14% of students who left a STEM undergraduate program cited its hostile and isolating atmosphere as a reason [158]. Since more women than men believe that competition is less conducive for their learning, self-efficacy, and achievement [85], the competitive nature of STEM may prompt women to drop STEM programs, more frequently than men [158]. We inspected the websites of the schools under study to understand whether the above observations applied to our dataset. We found that schools with a lower proportion of female applicants highlighted the competitions their teams had won in the past. These were often competitions related to sports or

STEM. For example, among the eight schools that produced a low proportion of female engineering applicants, the homepages of six contained photos of school teams that won recent competitions. On the other hand, the homepage of only one out of the eight schools that produced a high proportion of female engineering applicants displayed such pictures. Thus, a competitive environment and the pressure to win may have discouraged women from participating in these STEM activities, and ultimately from applying to engineering.

Observation #3: Both male and female applicants from schools with more female applicants report more personal influences and guidance from family members and friends. This can be seen in the applicants' response to why they are interested in engineering (Tables 5.5 and 5.6 in Section 5.3.3.2). This observation is consistent with past studies, which found that encouragement and guidance from parents and teachers was one of the most important factors why students, especially girls, chose STEM careers [143, 139, 16, 85, 220, 206]. Since parents and teachers greatly influence women's decision to pursue STEM careers [24], efforts should be made to reduce the STEM stereotypes they endorse and raise awareness about the variety of STEM careers. Various studies have confirmed that such efforts increase women's interest in STEM [24, 8, 9, 143].

Moreover, studies found that interactions with role models who did not endorse promale STEM stereotypes increased women's success beliefs in STEM careers [57]. Role model gender had no effect on these success beliefs [57]. Further, it was found that students who have access to role models, either in-person or in the form of videos or biographies, not only have an increased sense of compatibility with STEM, but are also more aware of STEM career possibilities [57, 287, 353]. This knowledge allows students to make informed choices about courses and career paths [57, 287, 353].

A possible explanation for having many accessible role models may be that applicants from schools with more female applicants belong to families with a higher socio-economic status or a background in STEM. Past research found that family socio-economic background and parental education attainments strongly predict the next generation's educational selection [327, 134, 31, 197, 133, 192, 220, 313, 191]. Researchers found that parents not only transfer cognitive and soft skills to their children (making them more employable), but they also provide encouragement, counselling, professional guidance, and resources [31, 134, 327, 133, 220, 206]. Students from families with a high socio-economic status report more parental conversations about college, more assistance with filling out college applications, and more peers planning to go to college [327, 220]. These factors affect students' clarity of career choice and preparedness for post-secondary degrees [327, 206, 220]. Therefore, providing all students with accessible role models, career counselling services, and guidance in college preparations can help level the playing field and increase the number of women interested in STEM careers.

Observation #4: We found that applicants from schools with high versus low proportions of female engineering applicants have different motivations to study engineering (see Engineering Interests and Goals and Reasons to apply to the university in Tables 5.5 and 5.6 in Section 5.3.3.2). Both male and female applicants from schools with more female applicants mention the love of science and their capability in STEM. Past studies found that these reasons are associated with applicants from families with high incomes and education levels [197]. Students from such families usually have access to more resources [327, 133, 228, 220], including high-quality science education and paid extracurricular activities. This may explain why applicants from schools with more female applicants mentioned more modern programming languages in comparison to applicants from schools with fewer female applicants (seen in Programming experience in Tables 5.5 and 5.6). Additionally, our observation is in line with studies that found that women were more likely to apply to college when they had access to higher quality resources, academic encouragement, and support from parents, teachers and peers [327, 220, 206].

On the other hand, applicants from schools with fewer female applicants appear to be interested in engineering for career security and growth as well as to contribute to society. Past studies found that these reasons are associated with applicants from families with low levels of education and income [197, 281, 192]. Valadez [327] argues that since students from low-income families are less likely to have access to resources, they pay more attention to (a) technology they can access through in-class learning and web sources and, (b) its implications on society. Therefore, a possible explanation for the aforementioned difference in motivations to pursue engineering may be a difference in the socio-economic backgrounds of the applicants [31, 134, 327, 192].

Observation #5: We found that applicants from schools that produce more female applicants report greater exposure and a wider breadth of interests in various subjects, including STEM. Applicants from these schools report learning more modern programming languages, more foreign travel, and a breadth of interests and experiences in fields such as music, sports, business, history, language, and community service. On the other hand, applicants from schools with fewer female applicants emphasize technical depth. This can be seen in the applicants' descriptions of their programming experience, extracurricular activities, jobs, and other information in Tables 5.5 and 5.6 in Section 5.3.3.2.

This observation may seem counter-intuitive since it seems to suggest that exposure to a variety of activities produces more female engineering applicants. In other words, to increase the number of female students who choose engineering, they should be exposed to various fields, not only STEM. This can be explained as follows. First, studies found that more females than males are highly skilled in both verbal and math domains [339, 181]. This may allow female students a greater variety of career options but also creates more ambiguous self-concepts and career goals [339, 181]. Participating in a variety of activities may have helped female students confirm their interest in STEM. Section 5.1 confirmed that female students who (eventually) applied to engineering programs reported a wider variety of interests than male applicants. Second, one study found that students who participated in a greater breadth of extracurricular activities received higher scores on the youth development index, including self-worth and psychological resilience [123]. Since STEM programs are male-dominated, it is more likely that women with firmer self-ability beliefs and resilience will choose them [337]. Third, a greater number of available options can increase the feeling of power and encourage (risky) decisions that otherwise would be avoided [210]. Participation in activities from various fields may have increased the number of career options available to women, empowering them to choose a traditionally male-dominated profession. Finally, a study found that an interest in creativity and design is a positive predictor of interest in computers and engineering [75]. Therefore, exploring non-STEM creative activities may also lead women to STEM.

Besides, exploring a variety of fields may not only increase the number of women who apply to engineering, but it may also help women be more certain about their interest in STEM and persist in the profession. A common reason stated by women leaving a STEM undergraduate program is their discovery of an aptitude for a non-STEM major that seems better suited for their interests, talents, personality, educational, career, and life goals [158]. Moreover, since more women prioritize fit, personal values, and lifestyle goals over interest while making career choices [339], participating in a breadth of activities may help them confirm personal and cultural fit before choosing a profession.

A possible explanation for having access to a breadth of activities and exposure may be a higher socio-economic background. As discussed earlier, parents with high economic status and education levels can afford to provide a variety of experiences to their children [31, 134, 327, 220, 228]. In addition, they may also understand the value of this breadth.

Other reasons may be school related. Upon inspection of school websites, we found that schools with fewer female engineering applicants had a more technical focus. These schools offered more STEM programs and highlighted participation in STEM competitions and extracurricular activities (organized by the school or local companies). On the other hand, schools with more female applicants appeared to offer a breadth of activities and programs in arts and technology. For instance, while almost all schools under consideration offered specialized programs and certifications in math, science, music, sports, and languages, schools that produced a high proportion of female engineering applicants offered additional programs in hospitality, transportation, health and wellness, biotechnology, geotechnology, aviation, non-profit, graphic design, business, culinary skills, string instruments, and other vocations.

There may be various reasons why fewer female students apply to engineering from schools with a greater technical focus. Studies found that emphasis on stereotypically male-dominated technical activities can make women feel discriminated against and trigger stereotype threat, decreasing their interest in STEM education and careers [302, 225]. Another study found that women's interest in STEM is reduced by exposure to stereotypical STEM objects and environments [56]. On a separate note, a study found that students who feel less control over their future are more likely to be less aware of what they want, and therefore fail to invest time and energy in pursuing it [204, 179]. Women were found to be twice as likely as men to agree to their parent's wishes to pursue STEM education and were, therefore, more likely to switch careers in the future as well [158]. Therefore, it is possible that schools' and parents' focus on technical education discourages women from choosing STEM undergraduate programs. Our findings suggest that women from backgrounds that offer a breadth of activities do not feel pressured to choose STEM and therefore are able to make an informed choice to pursue a STEM career.

5.3.5 Summary and Conclusions

In this section, we combined two datasets – undergraduate applications to engineering programs and high school statistics – to identify the characteristics of high schools that produce many female engineering applicants. We used statistical methods to identify correlations between the proportion of female engineering applicants from a given high school and the school's demographics and academic performance metrics. In addition, we used text mining methods on applicant responses to identify the unique characteristics of students from schools that produced many female engineering applicants. Unlike past work that focuses on female students' personal considerations to join engineering programs, our work contributes to the small yet growing body of work that studies the effect of students' surrounding systems, especially their high school contexts, on their choice of major.

We found that students from schools with more female engineering applicants reported more personal influences and guidance from family or friends, a breadth of interests, and interest and capability in STEM as reasons to study engineering as opposed to career security and growth. However, these students were less likely to participate in collaborative and competitive technical activities. Additionally, schools that produced more female engineering applicants reported a greater gender gap in math test scores, with female students performing better on average.

Our data-driven findings suggest that women still need an "extra push" towards engineering, in the form of role models or explicit evidence of being good enough at STEM.

Thus, increasing the availability of STEM role models during high school and nurturing female students' self-confidence in STEM subjects could help close the gender gap in engineering admissions. In addition, our results indicate that monitoring and reducing gender bias from female students' surroundings, particularly from STEM activities and competitions, may contribute towards this push. While this study focuses on ways of increasing applications from female students, a possible direction for future work is to repeat the analysis to investigate ways to increase applications from other underrepresented groups in engineering.

Chapter 6

Gender Differences in the co-op experiences of engineering students

Starting with the already low proportions of female students, with only 23% applying to and enrolling in engineering programs (Section 5.1), the proportion of women in the engineering workforce plummets to 17% [259]. To understand why this is the case, a lot of qualitative and quantitative work has been done on engineering students' career paths post-graduation. Studies have found work experiences, particularly dissatisfaction over pay and promotion opportunities, to be the primary reason behind female attrition [132, 157]. In addition, women have been found to leave the engineering work field due to discrimination, lack of social belonging, and a mismatch between the demands of their jobs and their lifestyle values of rearing a family and work-life balance [157, 317, 118, 339]. While most of these studies are based on later careers, some researchers argue that early career experiences may have a greater impact on subsequent career choices and drive attrition more than other factors [174, 132]. These researchers speculate that the gendered workplace experiences in later career stages, may in fact, stem from early career experiences. Since co-op jobs represent the first STEM work experience for many undergraduate STEM students, this chapter uses the Co-op dataset to investigate gender differences in early engineering careers.

Following Figures 1.1, 1.2, and 6.1, this chapter examines gender differences in every stage of the co-op pipeline. Starting with applications and concluding with work term evaluations, this chapter proceeds in the order shown in Figure 6.1, with each section dealing with a particular stage of the co-op process (shown in grey boxes). Additionally, Figure 6.1 provides an overview of the methods used to measure gender differences in each stage and the information it may provide, for example, a gender difference in choice, opportunity, perceived competency, and satisfaction (shown in white boxes). Overall, in this chapter, we explore gender differences in the jobs students apply to (Section 6.1), jobs they interview for and rank (Section 6.2), jobs they obtain (Section 6.3), their performance appraisals (Section 6.4), and their appraisals of their employers (Section 6.5). While most of the work discussed in this chapter has been published [66, 63, 65, 67], additional work on students' ranking strategies has been accepted but is pending publication [62].

This analysis of engineering students' early career experiences was enabled by access to the unique Co-op data extract (Section 3.3) containing work term information from applications to evaluations for 8,956 engineering students applying to 10,387 jobs (between September 2015 and August 2016). The number of enrolled students and the gender distribution of each program can be seen in Table 6.1. Since all these programs mandated co-op, all enrolled students are required to participate in the application stage of the co-op pipeline. As mentioned in Sections 3.3 and 4.2.1, in addition to analyzing gender differences in the co-op experiences of all engineering students (ENG), we also study gender differences in particular disciplines, for example, COMP and MECH (since these are the two largest disciplines in the dataset) and particular seniorities, for example, junior and senior students (definitions available in Section 3.3). Table 6.2 summarizes the sizes and the gender mix of the different populations under study. It also shows the gender distribution of students at various stages of the co-op process, namely, students who obtained an interview, were among the top-3 ranked choices of their interviewers, received an offer, or were placed (and hence evaluated).

By analyzing gender differences in the co-op experiences of these students, our work is the first in the field to study gender differences in co-operative education. In addition, since the Co-op dataset provides end-to-end information regarding a particular (co-op) labour market, our study is the first to examine gender differences as job candidates move along the various stages of applications, interviews, shortlists, acceptance, and evaluations. Not only are we able to analyze gender differences in internal employment decisions, for example, conversion from application to interview and shortlists, but the comprehensive, factual, and complete nature of the dataset allows us to investigate all job candidates competing for the same jobs. Overall, the goal of this chapter is to quantify gender differences in early career experiences and accordingly, provide data-driven actionable insight into closing the gender gap in STEM.

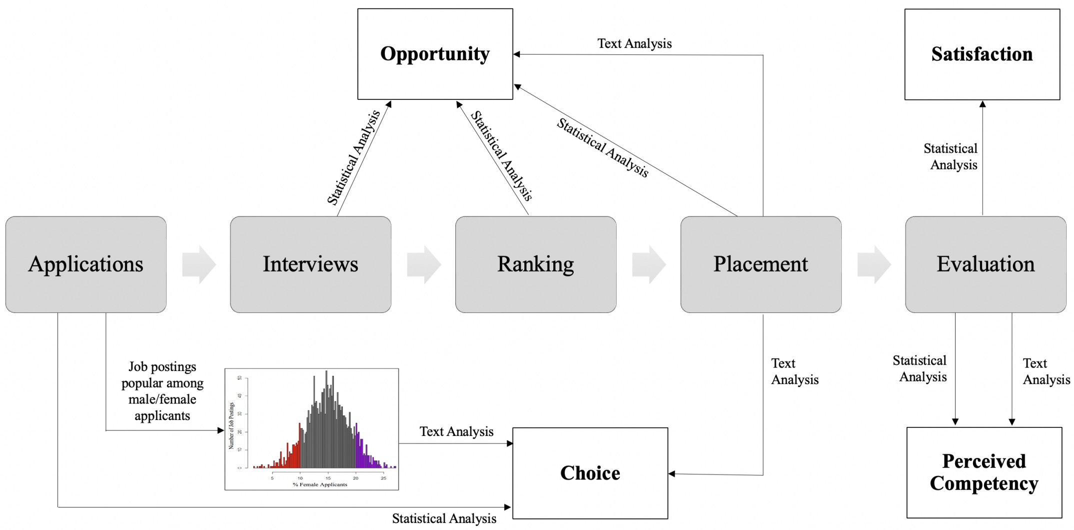

Figure 6.1: Methods used to analyze gender differences in the co-op pipeline

6.1 Gender differences in applications submitted

6.1.1 Motivation

To identify gender differences in students' co-op experiences, we start with the applications stage of the co-op pipeline. Past studies have identified differences in the kinds of careers chosen by men and women [278, 102]. Studies found that women prefer working with people, in socially oriented occupations, and have careers that benefit society, whereas men prefer working with things, for money and fame [338, 308, 278, 169]. Studies have identified this difference to be an important reason why fewer women pursue STEM programs and careers; they found that women may overlook engineering careers because they are considered incongruous with communal goals of collaboration and helping others [87, 88, 339, 85].

We analyze the application stage of the co-op pipeline to determine whether male and female students, already enrolled in engineering programs, apply to different kinds of co-op jobs (Figure 1.1). As students are free to apply to any job, differences in the number and job profiles that attract more male or female students may indicate a gender difference in choice or preference (shown in Figure 6.1). To the best of our knowledge, there is no previous work on identifying gender differences in preferences for jobs within the STEM field, especially during early careers.

Table 6.1: Gender breakdown by program

Program	Students	%Male	%Female
Mechanical	1108	88%	12%
Mechatronics	735	86%	14%
Computer	2724	84%	16%
Electrical	738	83%	17%
Software	657	83%	17%
Nanotechnology	409	75%	25%
Geological	151	69%	31%
Systems Design	465	67%	33%
Civil	614	66%	34%
Chemical	705	60%	40%
Management	306	58%	42%
Biomedical	85	44%	56%
Environmental	259	41%	59%
Total	8956	77%	23%

6.1.2 Data and Methods

The analysis is enabled by the Co-op dataset, which contains work term data for 8,956 engineering students applying to 10,387 co-op jobs in three semesters between September 2015 and August 2016 (refer to Section 3.3). We apply the statistical and text analysis methods mentioned in Sections 4.1 and 4.2 to determine whether male and female students apply to a different number and kind of co-op jobs (shown in Figures 6.1).

First, we calculate the number of applications sent by each student in a semester and use a t-test to compare the average number of applications submitted by male and female students (details about this statistical method can be found in Section 4.1). As mentioned in Section 4.2.1, we calculate this gender difference for all, junior, and senior co-op students enrolled in ENG, COMP and MECH (definitions presented in Section 3.3).

Next, we apply the term frequency analysis described in Section 4.2 to the job titles and descriptions of jobs that receive a much higher proportion of applications from male or female students in comparison to the other jobs within the discipline. Jobs in COMP and MECH are analyzed since jobs in ENG contain a mix of attributes from COMP and MECH, the two largest disciplines in the dataset. To do this, the method described in

Table 6.2: Gender breakdown during the different stages of the co-op process by job discipline

Group	Seniority	Students	Program / Applications %M	%F	Interviews %M	%F	Ranking %M	%F	Placement / Evaluation %M	%F
ENG	All	8956	77%	23%	78%	22%	77%	23%	77%	23%
	Junior	3828	74%	26%	74%	26%	73%	27%	74%	26%
	Senior	2144	81%	19%	81%	19%	81%	19%	81%	19%
COMP	All	3381	84%	16%	83%	17%	83%	17%	84%	16%
	Junior	1523	82%	18%	80%	20%	80%	20%	82%	18%
	Senior	693	87%	13%	87%	13%	85%	15%	87%	13%
MECH	All	1843	87%	13%	87%	13%	87%	13%	87%	13%
	Junior	780	83%	17%	82%	18%	81%	19%	83%	17%
	Senior	490	90%	10%	90%	10%	90%	10%	90%	10%

Section 4.2.1 is used to, first, identify the jobs corresponding to every discipline, and then, pick out those that receive a much higher proportion of applications from male students (labelled jobs_M) and female students (labelled jobs_F). Figure 4.1 of Section 4.2.1 shows an example of how jobs_M and jobs_F are identified within senior COMP jobs. Next, the parser with the *pre-processing* step specific to job titles and descriptions [59] (summarized in Section 4.2.2) is used to extract relevant tokens from all the job postings of jobs_M and jobs_F. Finally, the term frequency analysis described in Section 4.2.3 is used to identify (a) job attributes (i.e., tokens from job titles or descriptions) that frequently occurred in jobs_M and jobs_F, and (b) job attributes that occurred significantly more frequently in jobs_M or jobs_F.

6.1.3 Results

This section analyzes the gender differences in the application stage of the co-op process (Figure 1.2). The gender proportion of the students participating in the application stage (i.e., those who submit more than one application) can be found in Table 6.2. Following Figure 6.1, below we discuss the results of applying statistical and text analysis methods to the applications submitted. A gender difference in the applications submitted might reveal a gender difference in choice.

6.1.3.1 Statistical Analysis

Table 6.3 summarizes the results of applying statistical analysis on various stages of the co-op process, namely, applications, interviews, ranking, matching, and placements (refer to Section 4.1 for details of metrics). The table shows gender differences in all of ENG, just COMP, just MECH, and its junior and senior students. For each metric, if the difference between the male and female students of a group was statistically significant at a p-value of 0.05, the table reports the absolute difference, with M or F indicating whether the male or female outcome was higher. Differences that were not statistically significant at a p-value of 0.05 are marked by a hyphen (-). Additionally, differences that were statistically significant at a p-value of at least 0.05 are marked with *, at least 0.01 with **, and at least 0.001 with ***. Below, we discuss the gender differences in the metrics related to the applications stage.

As shown in Table 6.3, in ENG overall, there is no gender difference in the number of applications submitted. However, female COMP students, especially junior students, submitted slightly more applications than male students, and senior ENG male students submitted more applications than their female counterparts. However, we found no consistent pattern across the different groups of students.

6.1.3.2 Text Analysis

Looking into applications further, we conduct text analysis of jobs_{M} and jobs_{F}. Recall that jobs_{M} and jobs_{F} are job postings that received a much higher proportion of applications from male and female students respectively in comparison to the other job postings of the discipline. Below, we list (a) frequently occurring job attributes in jobs_{M} and jobs_{F} and (b) job attributes whose frequency is statistically significantly different in jobs_{M} and jobs_{F} (refer to Section 4.2 for explanation of methods used).

Frequent attributes: Tables 6.4a and 6.4b show the top 10 most frequent job title attributes in jobs_{M} and jobs_{F} of COMP and MECH respectively[1]. For example, 52% of jobs in COMP's jobs_{M} and 32% of jobs in COMP's jobs_{F} mention the job attribute "develop" (or its variants such as "developer" or "development", all reduced by the parser to "develop") at least once. The frequent job description attributes in jobs_{M} and jobs_{F} suggest similar trends and hence have not been shown.

[1] The results for all of ENG are omitted as they contain a mix of attributes from COMP and MECH, the two largest disciplines in the dataset. COMP's jobs_{M} and jobs_{F} contain 79 and 390 job postings, and MECH's jobs_{M} and jobs_{F} contain 178 and 461 job postings, respectively.

Table 6.3: Job application, interview, ranking, offer and placement statistics

Process	Metric	All			Junior			Senior		
		ENG	COMP	MECH	ENG	COMP	MECH	ENG	COMP	MECH
Applications	Avg # of applications submitted	-	F2.3**	-	-	F2.7**	-	M3.6***	-	-
Interviews	% students with $\geq$1 interview	-	F4.6%**	-	-	F7.8%**	-	-	-	-
	Avg # of interviews obtained	-	F0.7***	-	-	F0.9***	-	-	F1.1*	-
	Conversion rate	-	F1.4%*	-	-	F1.5%**	-	-	F5.0%*	-
Ranking	% students with top-3 rank	F2.7%*	F4.2%*	-	-	-	-	-	F9.0%*	-
	Avg # of top-3 ranks received	F0.2**	F0.3*	-	-	-	-	F0.3*	-	-
	Conversion rate (Interview to top-3 rank)	F4.9%***	-	F6.6%*	-	-	F9.7%*	F4.4%*	-	-
	% students with $\geq$1 offer	-	-	-	-	-	-	-	F14.3%**	-
	Avg # of offers received	-	-	-	-	-	-	-	F0.6*	-
	Conversion rate (Interview to Offer)	F1.6%*	-	-	-	-	-	F3.3%*	F6.0%*	-
Placements	% Employed Students	F1.1%*	-	-	-	-	-	-	-	-

Table 6.4: Top 10 frequent job title attributes in $jobs_M$ and $jobs_F$

(a) COMP

Token	$jobs_M$	Token	$jobs_F$
develop	52%	develop	32%
softwar	51%	analyst	18%
game	6%	softwar	16%
embed	6%	design	12%
system	5%	web	12%
applic	5%	qualiti	9%
mobil	4%	product	8%
analyst	4%	ui	7%
web	4%	qa	7%
manag	4%	ux	6%

(b) MECH

Token	$jobs_M$	Token	$jobs_F$
softwar	22%	assist	13%
mechan	20%	develop	10%
develop	16%	product	8%
design	15%	analyst	7%
embed	12%	project	5%
system	7%	design	5%
product	6%	softwar	5%
assist	6%	manufactur	4%
control	5%	manag	4%
robot	5%	qa	4%

In COMP, titles in both $jobs_M$ and $jobs_F$ suggest software developer positions, with some titles in $jobs_M$ indicating gaming and embedded systems, and some titles in $jobs_F$ mentioning user interfaces and experience (UI/UX) and quality assurance. Similar trends were seen in junior and senior COMP jobs (results omitted for brevity); notably, some senior $jobs_M$ titles were hardware-oriented whereas some senior $jobs_F$ titles suggested data science positions. Gender differences were also seen in MECH (Table 6.4b): some $jobs_M$ titles suggest mechanical and embedded systems positions, while some $jobs_F$ titles suggest more project management, analyst, and quality assurance roles.

Significant Differences: Table 6.5 shows the top 10 job attributes that are mentioned significantly more frequently in $jobs_M$ descriptions than in $jobs_F$ (on the left), and vice versa (on the right). The corresponding analysis of job title attributes reveals similar results and hence has been omitted. The lists are sorted by the difference in frequencies, abbreviated Δ, computed as the percentage of job postings mentioning an attribute in $jobs_M$ (or $jobs_F$) minus the percentage of job postings mentioning this attribute in $jobs_F$ (or $jobs_M$). As can be seen in Table 6.5a, in COMP, $jobs_M$ are more likely to mention programming terms and hardware, whereas $jobs_F$ include more mentions of clients and reporting. Similarly, in MECH, $jobs_F$ are more likely to mention project management skills (see Table 6.5b). These findings are in line with the frequent job title attributes seen in COMP's and MECH's $jobs_M$ and $jobs_F$ (seen in Tables 6.4a and 6.4b).

Next, we explore the significant differences in junior and senior $jobs_M$ and $jobs_F$ of COMP and MECH. Several differences are seen, starting with senior COMP $jobs_M$ having more hardware and embedded systems jobs than junior COMP $jobs_M$. Furthermore, senior

COMP jobs_F appear to shift to data analysis roles. In addition, senior MECH jobs_M appear to shift from manufacturing to design positions, while senior MECH jobs_F appear to shift from supporting and recording roles to project management (full results omitted for brevity). Again, these results are consistent with the frequent job title attributes of junior and senior COMP's and MECH's jobs_M and jobs_F mentioned earlier.

Overall, Tables 6.4 and 6.5, along with the discussion above, indicate a gender difference in choice. Apart from some common positions, male and female students, irrespective of seniority, apply to different kinds of jobs.

Table 6.5: Differences in frequency between job description attributes of COMP and MECH jobs_M and jobs_F

(a) COMP

Token	jobs_M	jobs_F	Δ
c++	37%	12%	25%***
linux	33%	14%	19%***
hardwar	28%	10%	18%***
c	27%	9%	18%***
algorithm	23%	6%	17%***
debug	24%	9%	15%***
framework	33%	18%	14%**
java	39%	26%	13%*
scale	22%	9%	13%***
github	15%	3%	13%***

Token	jobs_F	jobs_M	Δ
document	40%	14%	26%***
busi	47%	25%	22%***
css	27%	5%	22%***
client	32%	13%	20%***
report	31%	11%	20%***
html	31%	11%	19%***
process	42%	24%	18%**
focus	28%	10%	18%***
support	47%	29%	18%**
meet	27%	9%	18%***

(b) MECH

Token	jobs_M	jobs_F	Δ
hardwar	37%	8%	29%***
c	29%	6%	23%***
machin	30%	8%	23%***
system	70%	49%	21%***
softwar	59%	39%	20%***
c++	26%	7%	19%***
assembl	29%	9%	19%***
mandatori	31%	13%	19%***
embed	22%	4%	19%***
appli	50%	31%	19%***

Token	jobs_F	jobs_M	Δ
manag	50%	26%	24%***
assist	52%	31%	22%***
report	40%	20%	20%***
support	53%	33%	20%***
construct	25%	6%	19%***
document	38%	20%	18%***
servic	36%	18%	18%***
help	36%	20%	16%***
activ	38%	22%	15%***
projectmanag	29%	14%	15%***

6.1.4 Discussion

Our analysis of the application stage of the co-op process led to the following observation. While we found certain similarities, we also found some differences in the job profiles that attracted male and female students. For example, all students in computing applied to software, web developer, and analyst positions, and all mechanical students applied to software and design roles. However, in computing, more male students applied to jobs involving hardware, firmware, and embedded systems, whereas more female students applied to jobs involving user interface and data analysis. In mechanical and mechatronics, more male students applied to manufacturing jobs and more female students applied to project management positions. While past work has found gender differences in choices between STEM and non-STEM professions, this is the first work to find gendered preferences towards particular jobs within an engineering field.

A combination of reasons may explain why male and female students made different choices when applying to co-op jobs. Men and women have been found to have different goals [197] that influence their occupational orientations [278]. For example, women have shown more altruistic inclinations and a preference for people-oriented jobs [308] (seen in Section 5.1). Thus, a gender difference in career goals might have motivated the observed gender difference in choice of co-op jobs. Moreover, Section 5.1 found female students to have a wider variety of interests than male students; this difference in breadth of interests could have led women to focus on different types of STEM jobs. Raising awareness about the variety of available co-op opportunities and the different career possibilities in STEM, especially those with communal goals, might attract more female students to engineering [57, 287, 353, 338, 13, 221, 125, 207].

In addition to gender differences in career goals, a *gender difference in ability* may provide another explanation for our results. Wang & Degol [339] found that females were more likely than males to be highly skilled in both verbal and mathematical domains. Thus, perhaps female students apply to computing jobs that require both programming and user experience elements because they perceive themselves as having high technical and communication skills (besides being interested in these types of jobs). On the other hand, the finding could be a function of female students', either implicit or society-driven, low mathematical self-concept; their *competence beliefs* may have led them to apply to the less technical jobs of their field [309, 102, 339]. Interviews with co-op students may provide more insight behind this difference.

6.1.5 Summary and Conclusions

In order to investigate gender differences in early engineering careers, we use data extracts covering a year of co-op data from nearly 9,000 undergraduate engineering students enrolled in the co-op programs of a large university. While the data extracts contain information regarding all stages of the co-op pipeline, from application to evaluations, this section analyzes data from the application stage alone. We apply statistical and text analysis methods to determine whether male and female students send applications to different number and kinds of co-op jobs. To the best of our knowledge, this is the first work to find gender differences in preferences towards particular engineering jobs. In addition, since the analysis is based on comprehensive and real labour market data and is not restricted by the small and self-selected samples collected through surveys and interviews, it is the first study to examine gender differences in the application patterns of all competing job candidates.

We found that job profiles that attracted significantly more male or female students had some overlap as well as some differences. For example, while all students in computing applied to software development and analyst roles, female students in computing were more likely to apply to jobs involving user interfaces, user experience, and data analysis, whereas male students in computing were more likely to apply to jobs involving embedded systems, hardware, and firmware. Interviews with co-op students may provide more insight into the reasons behind this difference.

Since our finding suggests that male and female students enrolled in engineering programs have different preferences, highlighting the different types of available jobs may attract more female students to study engineering. For example, in addition to advertising co-op roles in software development and system analysts, students should be informed about co-op opportunities in user interfaces/user experience, data analysis/data science, and project management. Similarly, adding user experience and data analysis elements to curricula might also aid in aligning STEM's male-centric pedagogy with female students' goals and interests, thus attracting and retaining more female students in STEM [69]. Overall, our finding may be of interest to academic institutions and employers wishing to increase STEM enrolment and diversify the talent pool. In addition, it may also provide a starting point to investigate gender differences in preferences for courses and jobs within the different disciplines of engineering.

6.2 Gender differences in the Interview and Ranking stages

6.2.1 Motivation

This section analyzes gender differences in students' experiences during the Interview and Ranking stages of the co-op pipeline (Figure 6.1). Since studies that analyze later careers have identified gender differences in opportunity (specifically in terms of hiring, promotion, and pay) as the main reason why women leave engineering [157], the goal of this section is to determine whether male and female students receive different opportunities during early careers as well. In this section, we analyze students' co-op experiences to determine whether male and female students receive equal opportunity, especially in terms of the number of interviews and offers received.

In addition to the gender differences in the opportunities received, this section also analyzes students' responses to them. Recent work reports that competition related to interviewing for and securing co-op placements is a source of stress for students [253, 92]. Students who perceive a lot of competition and are not confident about finding a co-op job may accept any and all opportunities they receive as a way to maximize their chances of finding a job. Considering that early career experiences can greatly affect subsequent career choices [174], it is important that students from a particular gender are not more likely to be in such situations. Since gender differences in the opportunities received during co-op may lead to dissatisfaction, and in turn, attrition from engineering programs and careers, this section analyzes the number of opportunities received as well as how students respond to them. To the best of our knowledge, this is the first study to inspect gender differences in the internal employment decisions made by employers and students. Additionally, while findings of past studies are based on surveying particular employers and candidates, our observations are based on data from an entire (co-op) job market.

6.2.2 Data and Methods

The analysis is enabled by the Co-op dataset, which contains work term information from applications to evaluations for 8,956 engineering students (refer to Section 3.3 and Figure 1.2). For all, junior, and senior students in ENG, COMP, and MECH, this section measures gender differences in the Interview and Ranking stages of the co-op pipeline. Before moving on, let us recall the details of the Interview and Ranking stages. First, for each application submitted, students are notified about whether the employer shortlisted

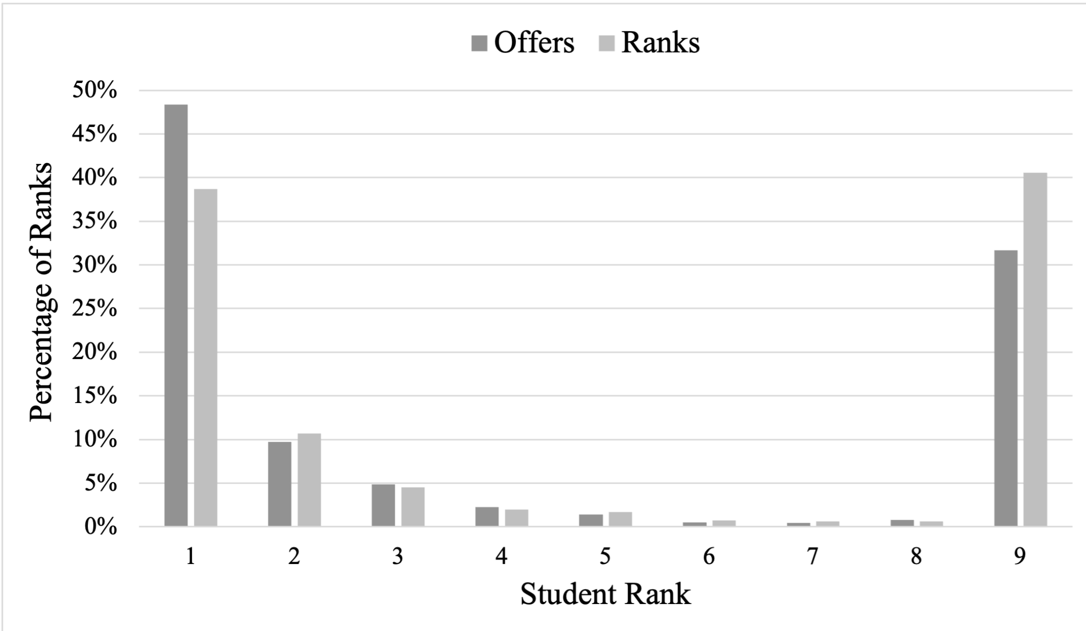

Figure 6.2: Distribution of student ranks

them for an interview. After the interview, a student is shown whether the corresponding employer made them an Offer, shortlisted them (Rank - but students are not shown the rank number), or is not willing to hire them (No Rank). The Interview and Ranking columns of Table 6.2 list the proportion of male and female students who received at least one interview or one Rank or Offer, respectively.

After employers rank students, students need to give a rank (between one and nine) to the employers who made them an Offer or shortlisted them. Ranking an Offer one means that the student is guaranteed to match with this job offer. However, ranking a Rank one may or may not lead to a match (it may lead to a match if the student who was offered this position does not rank the offer one). Since students are aware of this, they may react differently to the Offers and Ranks they receive. Figure 6.2 shows the distribution of ranks students give to Offers and Ranks. The most frequent student ranks include 1 and 9, with very few students ranking their options between 2 and 8. Additionally, Figure 6.2 shows that a higher proportion of students rank the Offers they receive one and the Ranks they receive nine. Finally, a matching algorithm that minimizes the sum of student-job rank pairs assigns students to jobs.

We apply the statistical analysis methods described in Section 4.1 to measure gender differences in the Interview stage. Specifically, we use the (1) proportion test to compare the fraction of male and female students who obtained at least one interview, (2) t-test to compare the average number of interviews obtained by male and female students, and (3) t-test to compare the average *conversion rate* of male and female students, which is the number of interviews obtained divided by the number of applications.

For the Ranking stage, we conduct the (1) proportion test to compare the fraction of male and female students who were a top-3-ranked choice of at least one interviewer, (2) t-test to compare the average number of top-3 ranks obtained by male and female students, and (3) t-test to compare the average interview to top-3 rank conversion rate, which is the number of top-3 ranks divided by the number of interviews. We repeat these tests with students who receive top-1 ranks, that is, Offers. In addition to considering students' Offers and Ranks separately, we also categorize them based on a combination of Offers (none, one, or more) and Ranks (yes or no) received. This combination defines a student's *situation* and may also determine their reaction. We use a proportion test to compare the percentage of female students in each situation with the proportion of female students who participate in the Interview stage.

As part of the Ranking stage, we also look into the gender differences in students' ranking decisions. Given that the matching algorithm is designed to minimize the sum of the ranks of the student-job assignments (refer to Section 1), students may use different ranking strategies depending on the level of competition they perceive. For example, students who do not receive any offers may be willing to take any job they were shortlisted for, and therefore, give the top rank of one to all (or multiple) employers in an attempt to maximize their chances of finding any job. On the other hand, more confident students in this situation may take a risk and indicate a preference for some employers over others by ranking their preferred options one, and other options nine (indicating that they have a strong preference against this option) or two, three, and so on. Furthermore, if a student strongly does not like any of their options and wishes instead to find a co-op job on their own (outside the institution's matching process), they may give all the jobs the lowest possible rank of nine. Similarly, students who receive offers may accept them or rank them lower depending on their confidence to obtain other opportunities. Since a student's ranking strategy is dictated by their assessment of current and future opportunities (i.e., the opportunities they currently have and the ones they believe they can obtain), it is important to identify gender differences in the different ranking strategies used by students.

To identify the ranking strategies used by students in different situations, we first identify the frequent *sets* of ranks they give to the Offers and Ranks they receive. For example, suppose a student receives three offers, of which they rank one one (i.e., they accept this offer), and give a rank of two to the two others; the set of ranks this student gives to the offers they receive is {1, 2}. Next, for every situation, we identify such common ranking patterns and group the ones with similar matching outcomes (referred to as ranking strategies). For example, student rank sets of {1, 2} and {1, 3} to the offers received are grouped together (shown in the second row of Table 6.8d) since they have the same matching outcome; they accept an Offer (and hence will be matched with it), but also indicate a backup choice. Finally, we use the two-proportion z-test described in Section 4.1 to compare the proportion of female students who use a particular ranking strategy versus those in the same situation. Gender differences in student situations and ranking strategies did not differ by discipline or seniority, and hence their results have been omitted for brevity.

Instead of the method described above, we tried various clustering algorithms, namely Agglomerative Hierarchical Clustering[2] and HDBSCAN[3], with multiple feature sets, distance metrics, scalers, and hyperparameters. However, since none of them divided the dataset into comprehensible clusters corresponding to the different ranking strategies used by students, the results of these methods are not reported in the thesis.

6.2.3 Results

This section presents the gender differences in the interview and ranking stages of the co-op process (Figure 1.2). We examine gender differences in (a) interviews and rankings received by students, and (b) students' responses to these rankings (referred to as students' ranking strategies). A gender difference in the above might indicate a gender difference in the opportunities received (Figure 6.1).

Interviews and Rankings received by students: Returning to Table 6.3, we look at the differences between the average number of interviews, offers, and top-Ranks received by male and female students. Several significant differences are noted in COMP, but not in MECH. First, female COMP students, especially junior students, are more likely to obtain interviews than their male counterparts. Second, senior female students in COMP are more likely to be top-3 ranked than male students. Third, senior COMP female students are more likely to receive offers. ENG and MECH demonstrate no gender differences in interview opportunities, but exhibit some differences in favour of female students in the number of ranks and offers received. The reasons behind the dissimilarities between COMP and MECH cannot be confirmed without further investigation. Overall, the statistical

[2] https://scikit-learn.org/stable/modules/generated/sklearn.cluster.AgglomerativeClustering.html

[3] https://hdbscan.readthedocs.io/en/latest/how_hdbscan_works.html

Table 6.6: Groups of students according to Offers and Ranks they receive

Label	Offer/s	Rank/s	%	%F
No Offer or Rank	No	No	19	20
Only Rank/s	No	Yes	29	24
Single Offer	Yes (1)	No	13	20
Single Offer & Rank/s	Yes (1)	Yes	16	22
Multiple Offers	Yes (>1)	No	5	21
Multiple Offers & Rank/s	Yes (>1)	Yes	17	25

analysis of the interview and ranking stages (results seen in Table 6.3) indicate that gender differences in opportunity, if present, exist in favour of female students.

To further understand gender differences in ranking, Table 6.6 groups students by the number of Offers and Ranks they receive. Every student who received at least one interview, is categorized into a situation depending on the combination of (a) Offers (none, one, or more), and (b) Ranks (yes or no) they receive. Table 6.6 lists the proportion of students and female students in each situation. In comparison to the percentage of female students among ENG students who receive an interview (seen in Table 6.2), Table 6.6 shows that a slightly higher proportion of female students receive multiple Offers and Ranks (last row in Table 6.6) and a slightly lower proportion of female students receive No Offers or Ranks (first row in Table 6.6). These findings align with the findings from Table 6.3 and suggest that in co-op settings, female students received more opportunities and were preferred over male students. COMP, MECH, junior, and senior students reveal similar gender differences and hereon, their results have been omitted for brevity.

Students' Ranking Strategies: Since students in different situations may react differently to the Offers and Ranks they receive, we separately analyze the ranking strategies for each student group in Table 6.6 and examine any differences in the strategies used by male and female students in the same situation. As mentioned before, students in a situation where they did not receive any Offers but were shortlisted may rank all their options one in order to increase their chances of finding any job. On the other hand, students who receive multiple Offers may rank the offers in order of preference. Since students who receive no Offers or Ranks do not rank any employers, they are excluded from further analysis.

For each group of students in Table 6.6, we identify their frequent ranking patterns, group ranking patterns that lead to similar matching outcomes (referred to as ranking strategies and shown in Table 6.8), and report the difference in fraction of female students who use different ranking strategies. As an example, Table 6.7 reports the frequent sets of ranks given to Offers and Ranks by students with multiple Offers and Rank/s. Most

Table 6.7: Most frequent sets of ranks given by students who receive multiple Offers and Ranks

To Offers	To Ranks	%
{1, 9}	{9}	45
{1, 2}	{2}	5
{1, 2}	{3}	3
{1, 2}	{9}	3
{1, 2, 9}	{9}	2
{1, 9}	{1, 9}	2
{1, 9}	{1}	1
{1, 9}	{2}	1
{9}	{9}	1
{1, 2, 3}	{4}	1

students in this situation accept one of their Offers (by ranking it one) and rank their other options nine (first row, capturing 45% of students). Other students accept an Offer and give ranks of two, three or nine to their other options. In rare cases, students give a rank of nine to all their options (second-last row). In this case, these students would most likely not be matched with any job.

We group ranking patterns into ranking strategies. Table 6.8e shows the ranking strategies of students who received multiple Offers and Ranks. The first row in this table corresponds to students who rank one of the Offers one (i.e., they are guaranteed to match with this job). The second row corresponds to students who, in addition to ranking an Offer one, also rank another option one. The third and fourth rows indicate the other ranking strategies that students use after accepting an Offer. Recall that once a student ranks an Offer one, they will be matched with it. Therefore, ranking strategies in rows two, three, and four indicate students' ranking preference without affecting their chances of finding a match. The fifth row corresponds to students who rank all their Offers greater than one (in turn, reducing their chance of finding a match). The sixth and seventh rows indicate how students rank other options after ranking all their Offers greater than one; the sixth row corresponds to students who take risks by ranking a Rank one instead of ranking an Offer one, and the seventh row corresponds to students who rank all their options greater than one (indicating that they do not consider any of the options to be ideal).

The second column in Table 6.8e, same as all tables in Table 6.8, shows the percentage of students with a particular ranking strategy as a proportion of the number of students in the situation. For example, Table 6.8e shows that among students who receive multiple

Table 6.8: Student ranking strategies

(a) Students who receive Only Rank/s

Label	%	%ΔF
Rank All 1	69	0
1 & Rank Any Other <9	8	-1
1 & Rank All Others 9	10	6*
Rank All >1	14	-4

(b) Students who receive a single Offer

Label	%	%ΔF
Rank Offer 1	95	-1
Rank Offer >1	5	9

(c) Students who receive single Offer and Rank/s

Label	%	%ΔF
Rank Offer 1	82	1
Rank Any Other 1	24	3
Rank Any Other <9	36	3
Rank All Others 9	40	-2
Rank Offer >1	18	-4
Rank Any Other 1	79	-3
Rank All Others >1	21	-5

(d) Students who receive multiple Offers

Label	%	%ΔF
Rank an Offer 1	98	-1
Rank Any Other <9	39	0
Rank All Others 9	61	-1
Rank All Offers >1	2	46**

(e) Students who receive multiple Offers and Rank/s

Label	%	%ΔF
Rank An Offer 1	93	0
Rank Any Other 1	9	-2
Rank Any Other <9	43	5*
Rank All Others 9	48	-4
Rank All Offers >1	7	-2
Rank Any Other 1	74	-1
Rank All Others >1	26	-3

Offers and Ranks, students either rank an Offer one (93%) or rank all of their Offers greater than one (7%). Among students who rank all Offers greater than one, 74% of students rank another option one (i.e., these students rank all Offers greater than one and rank a Rank one, in turn, taking a risk). The remaining 26% rank all their options, including the Offers received, greater than one (indicating that they do not consider any of the options ideal and want to reduce their chance of matching with any of them).

Finally, the third column labelled %ΔF reports the difference between the percentage of female students in a particular situation (those percentages are shown in Table 6.6) and the percentage of female students with a particular ranking strategy in that situation. If the difference is statistically significant at a p-value of 0.05, it is marked with an asterisk (*). Differences with a p-value less than 0.01 are marked with ** and differences with a p-value less than 0.001 are marked with ***. For example, Table 6.8e shows that, in comparison to the percentage of female students among students who receive multiple Offers and Ranks (25% according to the last row of Table 6.6), students who accept an Offer but still provide backup choices have 5% more (i.e., 30%) female students (third row, third column of Table 6.8e). The asterisk indicates that the proportion of female students in the two groups (25% vs. 30%) is statistically significantly different with a p-value less than 0.05. Tables 6.8a through 6.8d show the ranking strategies of students in other situations, along with the proportion of students and female students who use them.

Tables 6.8a through 6.8e show that students in different situations use different ranking strategies. The strategies appear to serve one of the following purposes: maximizing the chance of finding a match, reducing the chance of finding a match, indicating a preference even at the risk of reducing the chance of finding a match, or only communicating a preference. Below, we discuss gender differences in the ranking strategies used by students in various situations (Tables 6.8a to 6.8e).

First, we inspect student ranking strategies that maximize the chance of finding a match (first rows of Tables 6.8a to 6.8e). Students who have Offer/s can do this by simply accepting the Offer. However, students who are only shortlisted may use this ranking strategy because they are less confident in their ability to find a match. Tables 6.8a through 6.8e indicate that female students are not more (or less) likely to act this way.

Second, we examine students' ranking strategies that reduce their chance of finding a match (last row of Tables 6.8a to 6.8e). Students who use this strategy may not consider any of the available options ideal for them and be confident to find a co-op job on their own (outside the institution's matching process). Female students who receive Offer/s use this ranking strategy slightly more often (last row of Table 6.8b and 6.8d).

Third, we examine students who employ risky ranking strategies (second and third row of Table 6.8a and second-last row of Tables 6.8c and 6.8e). Students who use these strategies may perceive less competition and be more confident about finding a suitable match (within or outside the institution's matching process). The third row of Table 6.8a indicates that a higher proportion of female students who received no offers (i.e., who were only shortlisted) stated a clear preference for some of their options. This ranking strategy is risky as it reduces the chances of matching which is maximized if a rank of one is given to all the options.

Lastly, we discuss ranking strategies that only reveal students' preferences rather than their desire to maximize or minimize the chances of obtaining a particular job. This is done by analyzing how students rank their remaining options once they have accepted an Offer (the indented rows under the first row of Tables 6.8c, 6.8d, and 6.8e). Recall that if a student ranks an Offer one, they will get that job, regardless of how they rank their Ranks. Therefore, a possible explanation for using this ranking strategy could be to minimize the risk associated with a possible, yet rare, job cancellation. We found that a higher proportion of female students provided backup choices after accepting an Offer (third row in Tables 6.8c and 6.8e).

6.2.4 Discussion

The analysis of the interview and ranking stages of the co-op process was possible only due to the comprehensive nature of the Co-op dataset and led to two main observations. These observations revealed gender differences in the otherwise internal decisions made during an employment process.

Observation #1: Female students did not appear to be disadvantaged in the engineering co-op job search process in terms of interview and job opportunities; in fact, some metrics including the fraction of interviews that converted to offers and shortlists were in favour of female students (Table 6.3). On average, female students, especially those in computing, obtained more interviews and were shortlisted for jobs more often (and by more employers) than male students.

While some studies on post-graduate STEM employment found a hiring bias against women [107, 182, 232, 269, 285], our result is consistent with other work that discovered no bias or bias in favour of women [34, 49, 349]. Williams and Ceci [349] suggest that a pro-female bias could be due to anti-discrimination policies and other efforts to combat sexism in male-dominated workplaces. Other studies argue that women who enroll and persist in STEM degrees are more competent than average STEM men [339, 337, 144, 290]. Breda and Hillion [34] propose the "boomerang" effect as a possible explanation for the pro-female bias. They suggest that women who apply to highly skilled jobs do not elicit the general stereotypes regarding their motivation and ability; this induces a rational belief reversal in interviewers and increases their chances of being hired. Furthermore, they speculate that employers may have a conscious or unconscious preference for gender diversity, introducing a pro-female hiring bias in a male-dominated field [34]. A combination of these reasons may explain this observation.

Observation #2: The ranking behaviour of female students suggested that they were more inclined to take risks. Female students were not more likely to use a ranking strategy that would maximize their chances of finding a match and accept just any job. In fact, they were more likely to state a clear preference, and in turn take a risk, even when they were only shortlisted for jobs. Moreover, a higher proportion of female students who received only offers reduced their chance of finding a match by ranking all of them greater than one.

The use of riskier ranking strategies may suggest that female students perceived less competition and were more confident in their ability to find a co-op job, within or outside the institution's matching process. Past work suggests that risk-taking increases with the perception of power [210]. The availability of more co-op opportunities, as seen in observation 1, could have contributed to this perception. Furthermore, a study found that subjects who felt more confident saw more opportunities in a risky choice and therefore took more risks [186]. These feelings of self-confidence were found to be based on past success or feedback [186]. Similarly, in this study, female students' confidence and perception of (less) competition may have influenced their risk-taking capacity during ranking. This finding contradicts past work, which suggests that women, especially in male-dominated engineering fields, are less confident in their abilities and therefore are more likely to accept less desirable opportunities without stating any preferences [214, 280, 323].

In terms of ranking preferences, the study found that female students who accepted an offer stated their backup choices anyway. These students seem to be either overly cautious or risk-averse. Past work presented conflicting reports on whether women are more cautious or risk-averse than men [11, 41, 112, 238]. Thus, interviewing students about the competition they face in the co-op market may shed light on their intentions behind their ranking strategies and preferences.

6.2.5 Summary and Conclusions

In this section, we presented gender differences in students' experiences during the Interview and Ranking stages of the co-op pipeline. The analysis was facilitated by a dataset that recorded students' co-op experiences, from applications to evaluations, while they were enrolled in a co-op program of a large North American engineering institution. The goal was to determine whether engineering students faced any gender differences in opportunities during their early careers, especially in terms of the interview and job offers they received.

Unlike past studies that survey particular companies and potential employees to examine gender differences in opportunities, the Co-op dataset enabled us to study gender differences in the entire labour market. The study was based on comprehensive data of how potential employees navigate through a hiring process in a real (co-op) job market, giving insights into the gender differences in the otherwise internal screening processes adopted by both employers and students. To the best of our knowledge, this is the first work to study gender differences in (a) the co-op opportunities received by engineering students, in terms of both raw numbers and conversions from applications to interviews to offers, and (b) student strategies when making employment decisions and how these decisions may be impacted by students' perception of competition in the co-op job market.

Our analysis suggested that in a co-op environment with short work terms and a structured job search process maintained by the university, there is no evidence that female students are disadvantaged. In fact, female students appeared to obtain more interviews and offers than male students. Moreover, probably aware of this upper hand, we found female students to take more risks when ranking potential job opportunities. Since past research suggests that women's perception of STEM as "inhospitable male bastions" discourages them from pursuing STEM degrees and careers [349, 2, 302], female students interested in studying engineering may find these results encouraging. Therefore, an interesting direction for future work could be to set up an experiment that compares school students' perceptions of engineering programs before and after receiving the details of our observation.

6.3 Gender differences in Placements

6.3.1 Motivation

After analyzing the Application stage in Section 6.1 and the Interview and Ranking stages in Section 6.2, this section proceeds to inspect gender differences in co-op placements (refer to Figure 1.2). Since placements are related to where students applied to and what employment opportunities they received, analysis of placements may indicate a gender difference in both choice and opportunity (shown in Figure 6.1).

Past work has identified gender differences in the post-graduate employment of students with STEM degrees. For example, a study found female professionals with STEM degrees to be more likely than their male counterparts to work in the non-STEM fields of education and healthcare [15]. Studies have attributed this female attrition from the engineering workforce to primarily gender differences in the opportunities received, especially in terms of pay, promotion, and work profiles [157]. In this section, we examine the co-op work profiles of male and female students to understand if early engineering careers observe similar differences and whether these differences occur due to gender differences in choice or opportunity (Figure 1.1).

6.3.2 Data and Methods

The analysis is enabled by the Co-op dataset, which contains information about students who applied to, interviewed for, and got matched with a co-op job between September 2015 and August 2016 (refer to Section 3.3 and Figure 1.2). Following Figure 6.1, we apply statistical and text analysis methods to students' job placements to determine if (a) different proportions of male and female students were employed at the end of the job search process and (b) whether they had different types of work profiles.

We start by calculating the fraction of male and female students who participate in the co-op process and secure employment by the end of the matching process. We compare these fractions by using the two-proportion z-test described in Section 4.1. We present the results of this difference for all, junior, and senior co-op students in ENG, COMP and MECH (definitions mentioned in Section 3.3).

Then, to distinguish between the job profiles held by male and female students, we apply the term frequency analysis described in Section 4.2. First, we use the job description parser summarized in Section 4.2.2 to convert the job titles and descriptions of the jobs held by male and female students to tokens. Then, we use the term frequency analysis to compare the job attributes of the jobs held by male students with the attributes of the jobs held by female students (Section 4.2.3). For all students in COMP and MECH, we report two sets of attributes: (a) job attributes that frequently occurred in jobs filled by male and female students, and (b) job attributes that occurred statistically significantly more frequently in jobs filled by male or by female students. Since ENG contains a mix of COMP and MECH students, the two largest disciplines in the dataset, the jobs held by ENG students will contain a mix of attributes from COMP and MECH jobs and hence have not been studied.

As mentioned in Section 4.2.3, we drill into job placements further to determine if male and female students from different disciplines work in different types of co-op jobs. We apply the document clustering method described in Section 4.2.3 to first, identify the different types of co-op jobs held by students of each of the nine job disciplines (recall from Section 3.3 that the 13 academic engineering programs offered by the institution are mapped to nine job disciplines), and then, report the types of jobs in every discipline that contain a higher percentage of male or female students in comparison to the gender proportions of employed students in that discipline. To avoid drawing conclusions from small samples, we report only those types of jobs that employ more than 100 co-op students.

6.3.3 Results

This section presents the gender differences in the placement stage of the co-op process (the gender proportion of students who find a co-op job can be seen in Table 6.2). As seen in Figure 6.1, we use statistical and text analysis to examine gender differences in (a) the proportion of students who secure employment at the end of the matching process, and (b) attributes of the jobs filled. Since all prior stages, including applications, interviews, and ranking, affect matching and placement, a gender difference in placement might be a function of both choice and opportunity.

6.3.3.1 Statistical Analysis

Returning to Table 6.3, we see that the percentage of male and female students who were employed at the end of the job search process were the same in all groups except all of ENG, where 1% more female students were employed.

6.3.3.2 Text Analysis

To determine whether male and female students work in different kinds of co-op jobs, we apply term frequency analysis and document clustering to the text attributes of filled job postings (details in Section 4.2). Below, we report the frequent attributes of jobs held by male and female students (of COMP and MECH), followed by the word frequency differences between their postings.

Frequent attributes: Table 6.9 shows the ten most frequent job title attributes of jobs held by COMP and MECH male and female students; for example, the word "software" appears at least once in 50% of the job titles held by male COMP students and in 43% of the job titles held by female COMP students. The frequent job description attributes of jobs held by male and female students align with the frequent job title attributes (Table 6.9) and hence have not been shown. The frequent terms in all groups include technical terms such as "develop", "design", "software", "system" and "test", as well as references to soft skills such as communication and team(work). Looking at the frequent attributes of jobs held by male and female students, it appears that male and female students largely work in similar kinds of jobs. Junior and senior analyses (not shown for brevity) confirm this.

Significant Differences: The only attribute that appears significantly more frequently in the job titles of male COMP students than female COMP students is "software" (by 7%). On the other hand, job titles of the placements of female COMP students contain a variety of words that are mentioned more frequently (but with a Δ of less than 4%). These include tokens related to quality assurance, research, consultancy, and management from a variety of application domains including environment, health and trade. Similarly, "manufactur" is the only job title attribute more frequent in the placements of MECH male students, but those of female MECH students contain more references of "analyst", "projectmanag", and "research" in a variety of domains.

Table 6.10 shows the job description attributes of jobs held by male and female students, sorted by their difference of frequency (Δ) in the two groups. The job attributes of male COMP students include more programming and software development terms; for female COMP students, they include more business system analysis and data management terms. Jobs held by male MECH students contain more references to manufacturing jobs while those held by female MECH students contain more project management and software development terms. In all, the gender differences in job placements (Table 6.10) are similar to gender differences in job applications (Table 6.5), suggesting gender difference in placement opportunities to be a function of choice. Next, we present the results of document clustering, using which we identify the types of co-op jobs in the nine disciplines and the gender differences in them.

Table 6.9: Top 10 frequent attributes of job titles held by male and female students in COMP and MECH

(a) COMP

Token	Male	Token	Female
softwar	50%	develop	47%
develop	48%	softwar	43%
applic	7%	applic	8%
web	6%	web	7%
analyst	5%	analyst	6%
mobil	3%	qa	4%
test	3%	qualiti	4%
stack	3%	programm	4%
qa	3%	system	4%
assist	3%	mobil	3%

(b) MECH

Token	Male	Token	Female
mechan	15%	develop	13%
develop	14%	assist	13%
assist	11%	mechan	11%
design	11%	softwar	9%
softwar	11%	design	8%
manufactur	8%	product	6%
product	6%	project	4%
research	5%	manufactur	4%
system	3%	research	3%
project	3%	system	3%

Clustering: We start by analyzing the different types of jobs held by COMP students. As stated in Table 6.2, the male-female ratio of COMP students who found a co-op job was 84%-16%. Table 6.11 shows the eight largest clusters of COMP jobs, each row representing a particular kind of job (or cluster), sorted by size; the remaining three clusters were excluded since they had less than 100 jobs each. Each row includes a manually-assigned cluster label, the ten most representative words of the cluster centroid, the percentage of jobs in this cluster (out of all COMP job placements), and a percentage difference between female students in this cluster and all COMP placements (denoted by %ΔF). Differences in proportions that were statistically significant at a p-value of at least 0.05 are marked with *, at least 0.01 with **, and at least 0.001 with ***.

For example, the first row of Table 6.11 represents the largest cluster, containing 19% of all COMP job placements. The words in the cluster centroid suggest that the students of this cluster work in game development. The %ΔF of -1% indicates that the cluster has a female proportion of 15%, that is 1% less than 16% - the female proportion of COMP students who found a co-op job (shown in Table 6.2). As can be seen, this difference is negligible and is not statistically significant at a p-value of 0.05. Appendix A contains similar tables for co-op job placements of eight other disciplines, including MECH, Electrical, Nanotechnology, Civil, Industrial (contains co-op job placements of students from the academic programs of Systems Design and Management), Chemical, Environment (Environmental and Geological programs), and Biomedical. As explained in Section 3.3, we use the mapping provided by the institution to group academic programs into job disciplines.

Table 6.10: Differences in frequency between job description attributes of the placements of male and female students in COMP and MECH

(a) COMP

Token	Male	Female	Δ
featur	32%	26%	6%**
android	24%	18%	6%**
ios	19%	14%	5%**
improv	27%	23%	4%*
api	17%	13%	4%*
creativ	22%	18%	4%*
space	9%	5%	4%**
algorithm	17%	13%	4%*
store	9%	6%	3%*
hardwar	9%	6%	3%*

Token	Female	Male	Δ
document	31%	23%	7%***
busi	44%	37%	7%**
excel	39%	33%	6%**
communic	52%	46%	6%*
execut	18%	13%	5%***
sql	27%	21%	5%**
net	15%	10%	5%***
analysi	22%	17%	5%**
written	20%	16%	5%**
problemsolv	26%	22%	5%*

(b) MECH

Token	Male	Female	Δ
system	61%	48%	13%***
product	67%	56%	11%**
automot	17%	8%	10%***
test	50%	41%	8%*
tool	30%	22%	8%*
assembl	25%	16%	8%**
technic	41%	33%	8%*
manufactur	44%	37%	7%*
procedur	16%	9%	7%**
layout	13%	6%	7%**

Token	Female	Male	Δ
assess	15%	7%	8%***
written	22%	15%	7%**
problemsolv	27%	21%	6%*
client	18%	13%	6%*
creativ	15%	10%	5%*
consult	12%	7%	5%**
check	10%	5%	5%**
c#	11%	6%	5%**
profil	11%	6%	5%**
databas	13%	9%	5%*

Table 6.11: Largest clusters of COMP job placements

Label	Words in cluster centroid	%	%ΔF
Game Development	fun, passion, game, platform, java, agil, video, cloud, scalabl, web	19%	-1%
Development Jobs at Start-ups	awar, mandatori, python, featur, impact, scale, startup, javascript, stack, code	14%	-3%
Web Development	html, css, javascript, web, framework, jqueri, php, databas, mysql, sql	12%	-2%
Quality Assurance	script, qa, autom, defect, execut, methodolog, linux, agil, test, bug	10%	6%**
Application Development	android, ios, mobil, app, c, java, platform, c++, user, code	10%	0%
Business Systems Analyst	sql, c#, net, financi, server, invest, busi, databas, java, analyst	9%	3%
Support Analyst	troubleshoot, document, network, resolut, hardwar, enhanc, user, secur, resolv, support	9%	1%
Embedded Systems	c++, c, debug, hardwar, graphic, python, oop, linux, scienc, embed	9%	-5%*

Table 6.12 summarizes the types of co-op jobs in the remaining disciplines and categorizes them by their %ΔF. Instead of listing all cluster details (shown in Appendix A), for each discipline, the table shows the types of jobs with a negligible difference in gender proportions (i.e., %ΔF with a magnitude of <2%), a higher percentage of male students (i.e., negative %ΔF), and a higher percentage of female students (i.e., positive %ΔF). The asterisks indicate the strength of the statistical significance of the difference. Additionally, the table states the baseline proportion of female students in each discipline, that is, the percentage of female students in its placement stage against which the percentage of female students in each cluster is compared. Finally, the job disciplines are sorted by an increasing proportion of female students, with 13% female students in MECH and 57% female students in Biomedical.

Returning to COMP job placements, Table 6.11 indicates a negligible gender difference in game development, application development, and technical support jobs. However, there is a difference in the Embedded Systems cluster with 5% more male students working with the hardware and software of embedded devices (the proportion of female students in the Embedded Systems cluster is significantly lower, with its male-female ratio of 89%-11% compared to the baseline gender ratio of 84-16% in COMP placements). Moreover, the table suggests that more male students work as web developers or in development jobs that appear to be at technology startups, but these results are not statistically significant. On the other hand, a significantly higher proportion of female students work in quality assurance (with a male-female ratio of 90%-10% in the cluster).

Moving on to MECH jobs, Table 6.12 suggests that the only significant gender difference in co-op placements exists in the project management cluster, with 7% more female students than the baseline (the cluster has a male-female ratio of 80%-20% in comparison to the baseline ratio of 87%-13%). Overall, the gender differences found in the types of jobs held by COMP and MECH students align with the previous syntactic findings seen in Tables 6.9 and 6.10. Additionally, an inspection of the other disciplines suggests that, irrespective of the gender proportion of job disciplines, male and female students largely work in similar kinds of jobs with certain (small) differences (Table 6.12). While most clusters do not contain any significant gender differences, slightly more male students are involved in design and software development and slightly more female students are involved in project management and research.

Table 6.12: Labels of Clusters with more male or female students

Job Discipline	%F	Negligible Difference	Higher %ΔM	Higher %ΔF
MECH	13%	Hardware Design	Mechanical Design (2%), Manufacturing (4%)	Project Management (7%***), Web Development (2%)
Electrical	18%	Quality Assurance, Embedded Systems, Power Systems, System Development, Hardware Design	Application Development (4%)	Support Analyst (5%), Integrated Circuits (2%), Control Systems (3%)
Nanotechnology	26%		Software Development (3%), Manufacturing (2%), Electronics (7%)	Lab Assistant (8%), Optics (9%), Materials (4%)
Civil	34%	Traffic Planning	Capital Projects (5%), Structural Design (4%), Geotechnician (19%*), Estimator (17%)	Project Management (7%), Municipal Infrastructure (4%), Rail Inspector (8%), Diagnostics (4%)
Industrial	37%	Supply Chain Analyst, UI/UX	Quality Assurance (5%), Software Development (12%*)	Project Management (2%), Process Improvement (7%), Financial Analyst (5%)
Chemical	40%		Manufacturing (2%), Energy (16%**), Biotechnology (15%*)	Project Management (7%), Lab Assistant (9%), Nanotechnology (4%)
Environmental	49%	Survey and Design Jobs (arranged by student)	Civil Design (15%*), Hydrogeology (2%), Geologist (5%), Geotechnician (30%**)	Research Assistant (11%*), Water Infrastructure (7%)
Biomedical	57%		Systems Analyst (2%), Research Assistant (4%)	Clinical Analyst (15%), System Development (8%)

6.3.4 Discussion

Our analysis of the placement stage of the co-op process led to two main observations.

Observation #1: Female students did not appear to be disadvantaged in terms of co-op job opportunities; in fact, a slightly higher fraction of female students found co-op employment in comparison to male students (Table 6.3). This observation is in line with Section 6.2, which found that female students received more opportunities in the job search process than male students, particularly during the Interview and Ranking stages.

Related work on hiring practices in STEM careers present conflicting reports on gender bias. Section 6.2.4 discussed a combination of reasons that could have led to our observation regarding bias towards hiring female students. Broadly, Section 6.2.4 suggests that employers' conscious or unconscious pro-female ability bias or their preference for gender diversity may have led to this gender difference in opportunity. For details about past work that support these reasons, refer to Section 6.2.4.

Observation #2: Male and female students largely work in similar kinds of co-op jobs. While similar proportions of male and female students work in the technical jobs of their disciplines, slightly more male students are involved in design and software development and slightly more female students are involved in project management and research. These trends follow those seen in the applications submitted in Section 6.1 (elaborated below).

Past work found contradicting evidence regarding gender differences in STEM employment opportunities, especially in technical jobs or jobs requiring site work [232, 269, 303]. For example, some studies examined both math-intensive and non–math-intensive fields and found women to be preferred over identically qualified men when hiring teachers or university faculty [34, 349, 49]. Other studies found a bias towards men [232, 269]. Reuben et al. [269] found that, based on just looking at the candidate, both men and women subjects were twice more likely to hire a man than a woman for an arithmetic task. Similar biases towards men were found in faculty hiring [303] and the open source software website Github [316], where men's contributions were accepted more often than women's.

Gender differences in the kinds of co-op jobs held were very similar to the gender differences in the applications submitted, implying gender difference in co-op placement opportunities to be a function of choice. For example, in COMP, we found that both male and female students applied to and filled positions related to software development and analysis, but, jobs involving hardware, firmware, and embedded systems were applied to and filled by more male than female students. Similarly in MECH, more female students applied to and filled project management roles. Section 6.1.4 lists several reasons behind this gender difference in choice.

6.3.5 Summary and Conclusions

In this section, we presented gender differences in the co-op placement opportunities received by engineering students. The analysis was enabled by a dataset that recorded students' activities while they were enrolled in a co-op program at a large North American institution. The dataset contained information about the jobs students applied to, were shortlisted for, and the ones they finally matched with. In this section, we applied statistical analysis methods to determine if a higher fraction of male or female students found co-op employment, and text analysis methods to determine whether male and female students worked in different types of co-op jobs. To the best of our knowledge, this is the first work to find gender differences in the kinds of jobs filled by students of particular engineering majors, especially during early careers. In addition, since the analysis is based on a complete dataset with information ranging from applications to evaluations for all participating job candidates, the work is the first to provide insights into the differences in the number of placement opportunities received by male and female students competing in a closed (co-op) labour market.

In our analysis, male and female students appeared to be equally likely to secure co-op placements. In addition, they largely had similar kinds of job profiles, with certain exceptions. In all disciplines, slightly more male students were involved in design and software development, and slightly more female students were involved in project management and research. Furthermore, these gender differences in placements were similar to the gender differences observed during job applications (refer to Section 6.1), suggesting gender difference in co-op placement to be a function of choice rather than opportunity.

Since female students interested in joining engineering programs have cited male-centric stereotypes, including discrimination in hiring, to be a common reason why they avoid STEM programs [85, 339], our results might aid in changing their view of engineering and prompt them to reconsider. Additionally, incorporating elements preferred by female students (such as project management and research) into curriculum and job profiles may help institutions and co-op employers attract more female students.

6.4 Gender differences in the evaluations received

6.4.1 Motivation

This section analyzes the evaluations received by co-op students (Figure 1.2) and examines whether employers perceive male and female students to be equally proficient in the different aspects of their jobs (Figure 1.1). Studies on how STEM professionals are evaluated suggest gender differences in perceived competency. Not only are men (and their contributions) rated more highly than women [267, 96, 316], but they also tend to receive more actionable and task-oriented feedback in comparison to women who receive more critical and personality-related feedback [91, 291, 176, 47, 294, 77, 288, 38, 233]. Similar gender differences are found in teacher-student interactions in STEM classrooms, where teachers tend to attribute boys' success to ability and girls' success to hard work [320]. Research suggests that these gendered evaluations not only lower women's self-efficacy beliefs and performance in STEM, but they may also affect their career choices and opportunities [148, 339, 85, 222, 239, 288, 291].

Since identifying and eliminating gender bias from early career performance reviews can help retain more women in engineering programs and careers, this section focuses on identifying gender differences in both numeric and textual evaluations received by co-op students (Figure 6.1). To the best of our knowledge, our work is the first to study gender differences in early career performance evaluations of STEM students. In addition, our work is also the first to use text mining methods to analyze gender differences in performance reviews provided to STEM students or professionals.

6.4.2 Data and Methods

The analysis is enabled by the Co-op dataset, which contains records of students' and employers' activities during the different stages of the co-op pipeline (refer to Section 3.3). As can be seen in Figure 1.2, at the end of the four-month co-op placements, workplace supervisors and co-op students evaluate each other. Supervisors assess the performance of co-op students, where they rate them on various evaluation criteria and provide comments on their performance and further development. Following Figure 6.1, this section summarizes the statistical analysis methods used to measure gender differences in numeric evaluations, and text mining methods used to identify words that were used more frequently for male or female students.

To determine whether employers perceive any gender differences in the competencies of male and female students, we apply the following statistical tests to the performance evaluations received by students. First, we use the Mann-Whitney test to compare the average overall evaluation scores of male and female students. Then, for each of the 16 evaluation criteria listed in Table 3.1, we perform (1) a Mann-Whitney test to compare the average scores received by male and female students, (2) a proportion test to compare the fraction of male and female students who receive "Developing", "Good" and "Superior" scores, and (3) a proportion test to compare the fraction of male and female students receiving "N/A". We choose the Mann-Whitney test because of the ordinal nature of performance evaluations and present results for all, junior, and senior co-op students in ENG, COMP and MECH (definitions presented in Section 3.3). Definitions of the statistical tools used can be found in Section 4.1.

Next, we use the text mining methods described in Section 4.2 to understand gender differences in the written feedback and recommendations received by co-op students. Since the supervisor's comments are in free-text format, we use the parser described in Section 4.2.2 to extract words from the two text fields. In addition to the blank comments, comments that did not provide any meaningful feedback or recommendations were also removed during the *pre-processing* step of the parser (see Figure 4.2). These comments usually had a length of less than 30 characters. A few examples of such comments include, "Good luck!", "Discussed in-person", "None", and "Refer to above". Overall, a total of 5,708 feedback comments and 2,397 recommendations contained a meaningful evaluation and were converted to tokens.

Finally, for each text field (i.e., feedback and recommendations), we conduct a term frequency analysis to identify words that are used more frequently for male students than for female students, and vice versa (refer to Section 4.2.3). We report results for two groups of students: those from programs with less than 40% female students (the first nine in Table 6.1), and those from programs with greater than or equal to 40% female students (the last four in Table 6.1). Following Section 4.2.1, we initially analyzed the comments received by students from each discipline separately and observed that the comments received by students in the two groups mentioned above displayed similar trends. Thus, we omit per-discipline results for brevity. We also conducted the term frequency analysis on the comments received by students with different seniority levels and overall performance ratings[4]. The results followed similar trends and hence have not been shown in the report. To avoid overfitting, we ensured that each group being compared had more than 100 comments. Again, common English words with significant differences are excluded.

[4]With "Excellent" being the most common score received by students, they were divided into three categories: "Outstanding", "Excellent", and all students who were rated below "Excellent".

6.4.3 Results

This section presents gender differences in the end-of-term evaluations received by co-op students (Figure 1.2). The gender proportion of students who receive an evaluation can be found in Table 6.2. Following Figure 6.1, we apply statistical and text analysis to the numeric and textual evaluations received. A gender difference in the evaluations received might reveal a difference in how supervisors perceive male and female students and their competencies (Figure 1.1).

6.4.3.1 Statistical Analysis

Table 6.13 shows the gender differences in the overall performance rating and the 16 evaluation criteria for all, junior, and senior students in ENG, COMP, and MECH (using the Mann-Whitney test). The table follows the same format as Table 6.3: for statistically significant differences, an absolute difference in means is reported, with M or F to indicate whether the number was higher for male or female students; asterisks indicate the strength of statistical significance and hyphens indicate no statistically significant difference between means.

Table 6.14 provides more details about the gender differences in the 16 evaluation criteria for ENG students. Along with the differences in means (same as the first column of Table 6.13), Table 6.14 reports the results of proportion tests on comparing fractions of male and female students whose skills were rated as "Developing", "Good", "Superior" and "N/A" (recall Section 3.3). We omit the proportion test results for other groups in Table 6.13 as they showed similar trends as the Mann-Whitney results.

We start with the overall performance rating. According to Table 6.13, female students receive higher overall ratings in all of ENG and MECH, but there is no significant difference in COMP. Also, there is no significant difference between any group of senior male and female students.

We move on to the 16 evaluation criteria for all ENG students. Tables 6.13 and 6.14 show that in all of ENG, female students are rated more highly (and more likely to be rated "Superior") than male students on most criteria. Table 6.13 shows similar trends for both junior and senior female students. On the other hand, all and junior male students are rated more highly on resourcefulness and entrepreneurial orientation, but this trend does not persist in senior male students. Additionally, no difference is seen in ability to learn and problem-solving.

Table 6.13: Statistically significant differences between evaluation scores received by male and female students

Criteria	All			Junior			Senior		
	ENG	COMP	MECH	ENG	COMP	MECH	ENG	COMP	MECH
Interest in Work	F0.08*	-	-	F0.09*	-	-	-	-	-
Ability to Learn	-	-	F0.17*	-	-	-	-	-	-
Quality of Work	F0.12***	F0.12*	F0.16*	F0.14**	-	-	F0.13*	-	-
Quantity of Work	F0.13***	-	F0.17*	F0.16***	-	-	-	-	-
Problem-solving	-	-	F0.19*	-	-	-	-	-	-
Teamwork	F0.16***	F0.17***	F0.15*	F0.14***	-	-	F0.20***	F0.29**	-
Dependability	F0.15***	F0.14**	F0.17**	F0.15***	-	-	F0.16**	F0.21*	-
Response to Supervision	F0.10***	F0.16***	F0.13*	F0.10*	F0.12*	-	F0.14*	F0.21*	-
Reflection	F0.10**	F0.12*	F0.17**	-	-	-	F0.17**	F0.26*	F0.27*
Resourcefulness	M0.03*	-	-	M0.04*	-	-	-	-	-
Ethical Behaviour	F0.09**	F0.13*	-	-	-	-	F0.14**	-	-
Appreciation of Diversity	F0.11***	-	F0.16*	F0.10**	-	-	F0.15*	-	F0.30*
Entrepreneurial Orientation	M0.07**	M0.13**	-	M0.09**	M0.16**	-	-	M0.26*	-
Written Communication	F0.17***	F0.10*	F0.23***	F0.14***	-	-	F0.19***	-	F0.25*
Oral Communication	F0.09***	-	-	F0.07*	-	-	-	-	-
Interpersonal Communication	F0.17***	F0.12**	F0.25***	F0.15***	-	F0.19*	F0.23***	F0.28*	F0.35*
Overall Performance Rating	F0.08**	-	F0.19**	F0.12**	-	F0.27**	-	-	-

Table 6.14: Statistically significant differences between evaluation scores received by male and female students in ENG

Criteria	Mann-Whitney Test of Mean	Proportion Test of Developing (%)	Proportion Test of Good (%)	Proportion Test of Superior (%)	Proportion Test of N/A values (%)
Interest in Work	F0.08*	M0.49*	M2.66*	F3.17**	-
Ability to Learn	-	-	-	-	-
Quality of Work	F0.12***	M0.57*	M4.39***	F5.08***	-
Quantity of Work	F0.13***	M0.66**	M4.9***	F5.69***	-
Problem-solving	-	M0.84**	F2.59*	-	-
Teamwork	F0.16***	-	M5.8***	F7.02***	M0.87*
Dependability	F0.15***	-	M5.88***	F6.25***	-
Response to Supervision	F0.10***	-	M3.82***	F4.37***	-
Reflection	F0.10**	M0.51**	M3.45**	F4.45***	-
Resourcefulness	M0.03*	-	-	-	-
Ethical Behaviour	F0.09**	-	-	F5.42***	M3.62***
Appreciation of Diversity	F0.11***	-	M2.68*	F8.13***	M5.43***
Entrepreneurial Orientation	M0.07**	M0.66**	F2.46*	M3.21*	-
Written Communication	F0.17***	-	M7.42***	F8.21***	-
Oral Communication	F0.09***	-	M3.59**	F3.95**	-
Interpersonal Communication	F0.17***	M0.44*	M6.11***	F6.77***	M0.22*

Furthermore, Table 6.13 shows that the percentage of male (ENG) students who received "N/A" for teamwork, ethical behaviour, appreciation of diversity, and interpersonal communication, is significantly higher than the percentage of female students. This suggests that these qualities were either required, observed, or evaluated for fewer male students. Similar results were found in other programs.

Zooming in on COMP, Table 6.13 shows that all, junior, and senior male students are rated more highly than their female counterparts on entrepreneurial orientation, with other criteria either showing no difference (especially for junior female students) or some differences in favour of female students (especially for senior female students). Similar to COMP, female MECH students, especially senior students, are rated more highly than their male counterparts on several criteria. However, there is no significant gender difference in entrepreneurial orientation within MECH. Overall, with male and female students evaluated differently on many criteria, this analysis suggests a gender difference in perceived competency during early engineering careers.

6.4.3.2 Text Analysis

After analyzing numeric evaluations, we analyze gender differences in the textual feedback and recommendations received by students. As mentioned before, initially we analyzed the comments received by students from each program separately. However, we observed that the comments received by students from programs with less than 40% female students and programs with greater than or equal to 40% female students displayed similar trends. Thus, Section 6.4.3.2.1 presents word frequency differences in the feedback and recommendations received by male and female students enrolled in programs with $< 40\%$ female students and Section 6.4.3.2.2 presents the same for students from programs with $\geq 40\%$ female students. Since gender differences in the feedback and recommendations received by students with different overall performance ratings and seniority levels follow the same trends as the groups, we omit the details for brevity.

6.4.3.2.1 Gender Differences in Programs with $< 40\%$ Female Students

Feedback: Table 6.15 shows the differences in token frequencies in the feedback received by male and female students. On the left, Table 6.15 shows tokens that are mentioned statistically significantly more frequently in the feedback received by male students. On the right, it shows tokens mentioned significantly more frequently in the feedback received by female students. The lists are sorted by the difference in frequencies, abbreviated Δ, computed as the percentage of male (or female) students whose feedback mentioned a

Table 6.15: Word frequency differences in feedback received by male and female students enrolled in Programs with < 40% female students

Token	Male	Female	Δ
code	14%	10%	4%***
tool	7%	4%	3%**
fulltim	7%	5%	2%*
eager	2%	0%	2%*
written	3%	1%	2%*
prioriti	3%	1%	2%**
effici	2%	0%	2%*
hardwar	2%	1%	1%*
machin	2%	1%	1%*
analyz	1%	0%	1%*
expert	1%	0%	1%*
deadlin	1%	0%	1%*
iter	1%	0%	1%*
ecoop	1%	0%	1%*
tackl	3%	2%	1%*

Token	Female	Male	Δ
help	25%	20%	5%***
dedic	9%	5%	4%***
detail	7%	4%	3%***
collabor	5%	3%	2%**
thorough	6%	4%	2%**
enthusiast	5%	3%	2%**
addition	5%	3%	2%**
profici	3%	1%	2%***
delight	3%	1%	2%**
demand	2%	1%	1%**
timemanag	2%	1%	1%***
wonder	2%	1%	1%***
adapt	1%	0%	1%***
joy	1%	0%	1%**
potenti	1%	0%	1%***

token minus the percentage of female (or male) students whose feedback mentioned the same token. For example, feedback received by male students contained the word "code" 4% more often than feedback received by female students. Asterisks indicate the strength of the statistical significance of the difference, with all reported differences having a p-value of at least 0.05.

Feedback received by male students contains more technical terms. Table 6.15 shows that words relating to technical tasks, for example, "code", "tool", "written", "hardwar", "machin", and "analyz", are more frequent in the feedback received by male students. In addition, supervisors of male students are slightly more likely to refer to them as an "expert". This gender difference in the amount of technical feedback received exists in all groups with < 40% female students, irrespective of program, overall evaluation rating, or seniority.

Moreover, feedback received by male students contain more mentions of the word "eager" than feedback received by female students (see Table 6.15). Manual inspection of the comments containing the token "eager" reveals that these students suggest new ideas and take the initiative to start new tasks. In addition, as indicated by words such as "priori",

Table 6.16: Word frequency differences in recommendations received by male and female students enrolled in Programs with < 40% female students

Token	Male	Female	Δ
solut	8%	4%	4%**
seek	4%	1%	3%**
system	4%	1%	3%**
read	3%	1%	2%*
architectur	3%	1%	2%*
maintain	3%	1%	2%*
mistak	2%	0%	2%*
attent	3%	1%	2%*
web	1%	0%	1%*
algorithm	1%	0%	1%*
help	1%	0%	1%*
cooperat	1%	0%	1%*
opinion	1%	0%	1%*
hear	1%	0%	1%*
distract	1%	0%	1%*

Token	Female	Male	Δ
allow	8%	3%	5%***
express	4%	0%	4%**
network	4%	0%	4%***
oper	5%	1%	4%**
encourag	7%	5%	4%**
challeng	9%	5%	4%**
askquestion	9%	5%	4%**
general	4%	1%	3%**
varieti	3%	0%	3%***
afraid	3%	0%	3%**
shi	3%	0%	3%***
explor	4%	1%	3%*
market	3%	1%	2%***
tell	1%	0%	1%***
comfortzon	1%	0%	1%***

"effici", "deadlin", "iter", and "tackl" in Table 6.15, male students receive feedback on their efficiency and planning more often than female students.

Certain other words that occur slightly more frequently in the feedback received by male students include "fulltim" and "ecoop". Manual inspection of the comments containing the token "fulltim" indicates that the co-op employer has extended a full-time job offer to the student. The token "ecoop" refers to a program at the university that allows students to work in their own company (i.e., their start-up) for a co-op work term. Table 6.15 shows that the token "ecoop" is mentioned in the feedback of 1% male students and no female students.

Feedback received by female students, on the other hand, contain more references to their teamwork and interpersonal skills (indicated by words such as "help", "collabor", "delight", "wonder", and "joy" in Table 6.15). In addition, it contains more mentions of their thoroughness (indicated by words such as "detail" and "thorough" in Table 6.15), dedication ("dedic", "enthusiast"), and adaptability ("adapt").

In addition, tokens such as "addition", "potenti", and "demand" occur slightly more frequently in the feedback received by female students (seen in the table on the right in Table 6.15) and suggest that male and female students are referred to differently by their employers. Manual inspection of the comments containing the word “addition” indicates that female students are referred to as a “good addition to the team/company”. The word “potenti” is generally used in the phrase “has a lot of potential” and the word “demand” is used to describe a student’s ability to cope with a demanding work environment. These tokens are found more often in the feedback received by female students.

Recommendations: Table 6.16 follows the same format as Table 6.15 and shows the differences in token frequencies in the recommendations received by male and female students. Tokens in Table 6.16 suggest that male students receive more recommendations related to technical skills. This is suggested by words such as “solut” (stem of the word “solution”), “system”, “read”, “architectur”, “maintain”, “web”, and “algorithm”. In addition, male students are recommended to be more attentive to mistakes (indicated by the tokens “attent” and “mistak” in Table 6.16) and improve their teamwork and interpersonal skills (indicated by “seek”, “help”, “cooperat”, “opinion”, and “hear”).

On the other hand, female students are recommended to “express” themselves, “network”, not be “afraid” or “shy”, and ask more questions (see Table 6.16). Table 6.16 indicates that recommendations received by female students contained more occurrences of the words “allow”, “encourag”, “challeng”, and “comfortzon”. Manual inspection of comments containing these tokens suggests that female students were encouraged to challenge themselves and leave their comfort zones more often than male students. In addition, the recommendations received by female students contain more mentions of the tokens “oper”, “general”, “varieti”, “explor”, and “market” (see Table 6.16). Manual inspection of comments containing these tokens reveals that female students receive more recommendations to explore and increase their variety of knowledge, especially about business operations.

6.4.3.2.2 Gender Differences in Programs with $\geq 40\%$ Female Students

Tables 6.17 and 6.18 list the differences in word frequencies in the feedback and recommendations received by students enrolled in programs with $\geq 40\%$ female students. These tables follow the same format as Tables 6.15 and 6.16.

Feedback: Table 6.17 indicates that female students, in comparison to male students, receive more feedback related to their technical performance (suggested by tokens such as “applic”, “execut”, “user”, “technic”, “writtencomm”, “stack”, and “read”). Moreover, tokens such as “expertis” and “legaci” are found slightly more frequently in the feedback received by female students. On the other hand, feedback received by male students

Table 6.17: Word frequency differences in feedback received by male and female students enrolled in Programs with $\geq 40\%$ female students

Token	Male	Female	Δ
abil	22%	14%	8%**
understand	20%	12%	8%**
littlesupervis	9%	3%	6%***
effici	11%	6%	5%*
initi	7%	2%	5%**
pictur	4%	0%	4%*
surpris	5%	1%	4%**
devic	3%	0%	3%*
matur	3%	0%	3%*
prioriti	3%	0%	3%*
newtask	3%	0%	3%**
growth	2%	0%	2%*
difficulti	1%	0%	1%*
persist	1%	0%	1%*
ecoop	1%	0%	1%*

Token	Female	Male	Δ
hardwork	13%	6%	7%**
team	7%	3%	4%**
applic	6%	2%	4%**
execut	3%	0%	3%*
user	3%	0%	3%*
technic	3%	0%	3%**
comprehens	2%	0%	2%**
writtencomm	2%	0%	2%*
expertis	2%	0%	2%*
smart	1%	0%	1%*
stack	1%	0%	1%*
legaci	1%	0%	1%*
style	1%	0%	1%*
joy	1%	0%	1%*
read	1%	0%	1%*

references their "ability". This is in contrast to the results presented in Section 6.4.3.2.1, where male students received more technical feedback than female students.

Nevertheless, some of the gender differences in the feedback received by students of this group are similar to those found in programs with $< 40\%$ female students (Section 6.4.3.2.1). For example, even among students from programs with $\geq 40\%$ female students, male students are more likely to receive feedback on their eagerness to start new tasks (suggested by the tokens "newtask" and "initi" in Table 6.17, where "initi" is the word stem for "initiate" and "initiative"). They are also more likely to receive feedback on their planning and efficiency ("effic", "pictur", "prioriti"). The token "littlesupervis" in Table 6.17 indicates that supervisors find male students to be more independent than female students. Moreover, the token "ecoop" is mentioned in the feedback of 1% male students and no female students. Similarly, female students of this group are more likely to receive feedback on their hard work, thoroughness ("comprehens", which is the word stem for "comprehensive"), teamwork, and interpersonal skills, in comparison to their male counterparts (Table 6.17).

Table 6.17 shows that feedback given to male students contains more mentions of the

Table 6.18: Word frequency differences in recommendations received by male and female students enrolled in Programs with $\geq 40\%$ female students

Token	Male	Female	Δ
say	4%	0%	4%*
mistak	3%	0%	3%*
reserv	3%	0%	3%*
team	3%	0%	3%*
public	3%	0%	3%*
speak	3%	0%	3%*
open	3%	0%	3%*
expect	2%	0%	2%*
distract	2%	0%	2%*
error	2%	0%	2%*
topic	2%	0%	2%*
softskil	2%	0%	2%*
listen	2%	0%	2%*
respect	2%	0%	2%*
complex	2%	0%	2%**

Token	Female	Male	Δ
oper	5%	1%	4%*
creativ	5%	1%	4%*
surround	4%	0%	4%*
knowledg	4%	0%	4%*
instinct	3%	0%	3%*
quick	3%	0%	3%*
generat	3%	0%	3%*
difficult	3%	0%	3%*
system	3%	0%	3%*
learn	3%	1%	2%***
document	2%	0%	2%*
explor	2%	0%	2%*
interest	1%	0%	1%*
compani	1%	0%	1%**
deal	1%	0%	1%***

words "surpris", "growth", "persist", "difficulti", and "matur". Manual inspection of comments containing these terms revealed that these employers were pleasantly surprised to see the students' growth, persistence, and maturity.

Recommendations: Table 6.18 indicates that male students are referred to as "reserved" and are recommended to "speak" more often than female students (suggested by tokens such as "reserv", "say", "public", "speak", and "open"). This is in contrast to the results reported in the previous section (Section 6.4.3.2.1), where female students were recommended to ask more questions.

Table 6.18 also indicates that female students receive more technical recommendations than male students. Tokens such as "creativ", "knowledg", "generate", "system", "interest", "document", and "learn", occur more frequently in the recommendations received by female students. On the other hand, recommendations received by male students contain more occurrences of the tokens such as "topic" and "complex". Again, this is in contrast to the results shown in Section 6.4.3.2.1, where male students received more technical recommendations.

Nevertheless, some recommendations given to male and female students in programs with $\geq 40\%$ female students are similar to those in programs with $< 40\%$ female students (Section 6.4.3.2.1). For example, similar to male students from programs with $< 40\%$ female students (Section 6.4.3.2.1), male students from programs with $\geq 40\%$ female students are also recommended to keep an eye out for mistakes (indicated by "mistak", "distract", "error" in Table 6.18) and improve their teamwork and interpersonal skills ("team", "soft-skill", "listen", "respect"), more often than their female counterparts. On the other hand, female students from programs with $\geq 40\%$ female students are recommended to gain operational knowledge just like female students from programs with $< 40\%$ female students (indicated by words including "oper", "surround", "explor", and "compani" in Table 6.18 and confirmed by manual inspection of comments containing these tokens).

6.4.4 Discussion

The analysis of both numeric and textual evaluations received by co-op students reveals differences in how workplace supervisors perceive male and female students.

Observation #1: Statistical analysis of evaluation scores found that female students were rated equally or more highly than male students, except on specific criteria such as entrepreneurship orientation. Female students scored higher than their male counterparts on communal qualities (indicated by criteria including, teamwork, interpersonal communication, and appreciation of diversity), thoroughness (quality of work and reflection), adaptability (response to supervision), dedication (dependability), and written communication. Gender differences in either students' abilities or employers' perceptions of students' competencies might have led to this finding.

Gender differences in ability: A possible explanation for the above finding is that women who decide to pursue male-dominated degrees are likely to be highly qualified. For example, one study found that more men than women with low high school mathematics scores pursue STEM degrees [144]. Specifically, we found that female students tend to be evaluated more highly on written, oral, and interpersonal communication. Wang & Degol [339] found that girls are more likely to possess both high mathematical and verbal abilities, and boys are more likely to demonstrate higher mathematical abilities relative to their verbal abilities. Furthermore, we found that female students receive higher evaluation scores on teamwork. A recent report similarly found that girls outperform boys in collaborative problem-solving in several countries [241]. With the growing awareness of the importance of collaborative efforts, even in traditionally competitive fields such as STEM [28], this difference warrants further investigation.

On the other hand, male students in computing were perceived as having an entrepreneurial orientation more often than female students. Past work on risk-taking ability presented conflicting reports on how risk-averse men and women are [238]. Given the importance of entrepreneurship in today's economy, it is important to identify why a particular gender group receives higher evaluations in this area.

Gender differences in perceived competencies: In terms of evaluations on specific criteria, the evaluator's conscious or unconscious stereotypes about men and women could have influenced their evaluations. Studies examining gender differences in performance evaluations of professionals from various fields, including technology, military, politics, medicine, and law found that women were rated higher than men on communal qualities (e.g., those related to social relationships), while men were rated higher on agentic qualities (e.g., those related to goal achievement) [148, 91, 291, 176, 47, 294, 77, 288]. Particularly in STEM, the "boomerang" effect (explained in Section 6.2.4) could have created a pro-female bias in supervisors [34], making them evaluate women differently than men.

The analysis of the comments received by students led to three additional observations.

Observation #2: We found the following gender differences in the comments received by all groups of students, irrespective of their overall performance rating, seniority, and gender composition of their academic programs.

1. Female students are more likely than male students to be appreciated for their thoroughness, dedication, hard work, adaptability, teamwork, and interpersonal skills.

2. Male students are more likely than female students to be appreciated for their eagerness, planning, efficiency, and independence.

3. Female students are recommended to increase their business knowledge, including general information about the market and company operations.

4. Male students are recommended to keep an eye out for mistakes and improve their teamwork and interpersonal skills.

These findings align with the gender differences observed in numeric evaluations (Observation #1). For example, entrepreneurial orientation, the specific numeric criterion on which male students are rated more highly than female students, is closely tied to independence. Similarly, thoroughness is reflected in multiple numeric criteria including, quality of work, dependability, and reflection, all of which indicate higher scores for female students.

Similar to gender differences in numeric evaluations, these gender differences in feedback and recommendations may be due to gender differences in (a) how employers perceive their students' competencies, (b) opportunity, or (c) students' abilities.

Gender differences in perceived competencies: The gender differences we found are consistent with past studies, which examined feedback in education and in the workplace. In line with our findings and as mentioned before, studies examining professionals from various fields found that women were appreciated for their communal qualities (e.g., those related to social relationships) and men were appreciated for their agentic qualities (e.g., those related to goal achievement) [91, 291, 176, 47, 294, 77, 288]. In addition, women were tagged as "enthusiastic", "organized", and "unaware" and men as "analytical", "dependable", and "irresponsible" [291]. Studies in STEM classrooms found similar differences, with teachers associating boys with ability, and girls with hard work [320]. Social scientists and psychologists confirm the existence and prevalence of these gender stereotypes [148, 200], thus, suggesting the unconscious gender bias of the evaluator (i.e., the work term supervisor) as a possible reason behind the gender differences we found.

Studies suggest that the positive and negative gender stereotypes found in evaluations affect students' self-image and career choices [148, 339, 85, 222, 239]. Additionally, experiments found that gendered language in performance evaluations may affect hiring and promotion decisions [288, 291]. For example, when conducting a blind review of candidates for promotion, participants were more inclined to choose candidates described as "good at taking initiative". Since these (agentic) characteristics occur in the performance evaluations of men more often than women, this may lead to fewer promotion opportunities for women. Additionally, the study noted that participants considered collaborative skills, and thus unknowingly, women's profiles, less suitable for leadership roles [288]. Overall, since task-oriented qualities are more valuable to an organization than social-oriented qualities [72], the gender stereotypes in performance evaluations may give men a better chance to be hired, promoted, and more highly paid.

Thus, gendered evaluations may not only affect students' self-image, but it may also contribute towards lower pay and promotion opportunities for women, which is one of the leading causes for their attrition from STEM [157]. Therefore, eliminating gender bias from performance reviews, especially during early career stages, can help plug the "leaky" pipeline. To do so, universities offering co-op programs could communicate with participating co-op employers and emphasize the importance of unbiased feedback. Additionally, since the problem with implicit bias is that many people are not aware that they are biased, ensuring diversity training for workplace supervisors may help them identify and change their implicit perceptions.

Gender differences in opportunity: Gender differences in available opportunities could have led to the differences in the competencies found. We found that female students were appreciated for their adaptability more often than male students, indicating that perhaps female students were initially perceived to be more incompatible with the company culture. Past studies suggest that the masculine work and after-work culture of male-dominated professions make women uncomfortable [285]. This masculine culture may cause female students to consciously or unconsciously limit their workplace interactions (with peers and supervisors), limiting their access to operational knowledge. Given that the engineering fields have fewer female supervisors [259], female students may have found it difficult to communicate within a male-dominated hierarchy.

Gender differences in ability: Biological or society-driven differences in ability may have led to the gender differences in performance evaluations reported in this study. In addition to demonstrating superior verbal qualities in comparison to men [339], women have been found to possess more teamwork skills and altruistic tendencies [241, 169, 124, 88]. In fact, past studies have found women to be more interested in people-oriented activities and professions [308, 1, 68, 278, 68]. Thus, the gender differences in the competencies found could have been an artefact of this gender difference in ability and interest. Our analysis of numeric evaluations agrees with this finding and shows that female students score higher on teamwork and interpersonal communication (Section 6.4.3.1). In addition, our analysis of gender differences in engineering applicants (Section 5.1) found women to be more altruistic than men.

Overall, our results suggest that male and female students are perceived differently in the STEM workplace from the beginning of their careers. Whether these gender differences are due to employer perceptions or differences in opportunities or competencies cannot be determined directly from our data. However, regardless of the underlying reasons, we argue that universities offering co-op programs should communicate with participating employers to emphasize the importance of unbiased feedback in talent recruitment and retention.

Observation #3: There appears to be a relationship between the gender composition of academic programs and the comments received by students in those programs. We found that in programs with $< 40\%$ female students, a higher proportion of male students received feedback on their technical performance in comparison to female students. The recommendations received by male students also contained more technical directions for improvement. On the other hand, female students were recommended to participate, be less shy, and ask more questions. For programs with $\geq 40\%$ female students, the opposite is true. In these programs, female students receive more technical feedback and recommendations, and male students are recommended to be less reserved and speak more openly. This observation is particularly noteworthy because it occurs in a field with (traditionally) pro-male ability beliefs. The trend exists across all groups of students, irrespective of overall performance scores and seniority.

Gender differences in technical evaluation: The above observation is consistent with past observational studies that analyzed gender differences in teacher-student interaction and the feedback received by secondary school students. Some studies found that boys received more attention and feedback, particularly praise, criticism, and technical information, irrespective of the subject being taught (sports, modern languages, mathematics, science, and humanities) [239, 240, 93]. However, this was reversed in classes that contained as many or more girls [93]. Since feedback and recommendations on technical skills are important for all co-op students [72], universities may want to ensure that co-op evaluation forms include explicit requests to comment on students' technical skills.

Past studies that analyzed the performance reviews of men and women in various contexts, including technology and professional-services firms [288], a leadership development program [91], and the navy [291], found more mentions of technical words in the feedback received by men than women. These gender differences in technical feedback were attributed to the pro-male ability bias that exists in these fields. However, since all of these studies investigated samples containing less than 25% women, our results suggest the need for further investigation.

Gender differences in participation: A study conducted in a secondary school reported that both boys and girls participated more when their own gender was the majority gender in the classroom [93]. This was found irrespective of the subject being taught. Similarly, a study where engineering students were randomly assigned to teams (or "micro-environments") with varying gender composition reported similar conclusions. This study found that when women were the minority in a team (less than 25%), they spoke less, were less involved in teamwork, and felt less confident than women assigned to teams where they were in the majority (75% or more) [84]. This was true regardless of the students' academic seniority. Moreover, women from male-majority teams reported lowered engineering career aspirations after the team interaction [84].

Past studies attribute the reason behind this difference in participation to isolation (or social-belongingness concerns) and stereotype threat (the concern that one will be judged in terms of a stereotype) [93, 84]. Women were more affected by the gender composition in a classroom, leading to recommendations to create single-sex or gender-parity micro-environments (e.g., in-class teams or study groups) [84, 93]. Researchers experimenting with varying proportions of men and women in engineering teams found that gender-balanced micro-environments are particularly important for first-year students, to ensure that these students do not lose confidence and drop out of STEM fields [84]. Gender-balanced micro-environments helped students focus on learning, participate more freely, and in turn, gain the confidence to persist in gender-imbalanced environments. Another study found that participation in social-belonging interventions during student orientation programs improved women's social attitude and academic performance in male-dominated STEM programs [335].

Our results similarly suggest that co-op students working in environments where they are not visibly in parity participate less in team activities and may need additional encouragement. Following past studies, gender-imbalanced classrooms and workplaces could experiment with social-belonging interventions and gender-parity micro-environments and note their effect on student confidence.

Observation #4: Different words were used to describe male and female students in programs with $< 40\%$ female students and programs with $\geq 40\%$ female students. Phrases including "has a lot of potential", "challenge yourself", "allow yourself to grow", and "come out of your comfort zone", are more common in the comments received by female students from programs with $< 40\%$ female students. On the other hand, phrases including "surprised by performance" and "mature" are more common in the comments received by male students from programs with $\geq 40\%$ female students.

Studies of tokenism support the above observation and suggest that bias against a group occurs when said group is a minority in *any* given field [173]. Related work on minority groups (in terms of race and gender) presents conflicting reports on whether the feedback provided to those groups is more lenient or harsh [47, 294, 145]. Most studies that report gender differences in feedback note that the same trait is described more positively for men than for women [91, 47, 294, 77, 38]. However, all these studies were conducted in male-dominated professions.

6.4.5 Summary and Conclusions

In this section, we analyzed gender differences in early career workplace performance reviews. To do so, we used a unique dataset containing work term evaluations of students enrolled in undergraduate engineering co-op programs. We used statistical tools to analyze gender differences in the numeric evaluations received by students and text analysis to analyze word frequency differences in employer feedback and recommendations for professional development. Since gender differences in perceived competency, especially during early careers, can lead to dissatisfaction and attrition, early identification and elimination of these differences can promote female students towards engineering programs and careers. While past work has analyzed gender differences in post-graduate employment of engineering professionals, this is the first work that analyzes gender differences in the evaluations received by engineering students during their early careers.

The numeric performance appraisal analysis revealed that female students were rated equally or more highly than male students on most criteria, except specific criteria such as entrepreneurship. These results may be used to combat stereotypes regarding gender difference in STEM ability [309, 169]. Furthermore, universities and employers may want to provide resources to help students acquire the skills they were rated low on. With the growing importance of soft skills, including communication, collaboration, and entrepreneurship in engineering professions [71], minimizing the gender gap in these skills can help maintain a diverse workforce.

Text analysis of supervisor comments revealed that male and female students were perceived differently in the STEM workplace from the beginning of their careers. We found that male students were appreciated for taking initiative more often than female students. In addition, they were described as efficient and independent and were recommended to improve their interpersonal and teamwork skills. On the other hand, female students were appreciated for being thorough, hardworking, social, and collaborative and were advised to gain business knowledge more often than male students. Furthermore, we found a possible link between the gender composition of the programs and the comments received by the students. While the visible majority gender was more likely to receive technical feedback and recommendations, the visible minority gender was advised to work on their confidence and ask more questions.

Since reiteration of gendered feedback may lead to career dissatisfaction and attrition [222, 288, 157], our results emphasize the importance of unbiased feedback in early career settings such as co-op work terms and internships. Universities offering co-op programs should communicate with participating employers to emphasize the importance of unbiased feedback in talent recruitment and retention and at the same time, provide resources to help students acquire certain skills. Moreover, special attention should be paid to encourage minority groups. An interesting direction for future work is to interview STEM alumni to determine if their co-op experiences, specifically the feedback they received, affected their career paths. It may also be useful to investigate the effect of the workplace supervisor's gender on performance reviews (we were unable to do this analysis because our dataset did not include any information about workplace supervisors).

6.5 Gender differences in satisfaction

6.5.1 Motivation

This section analyzes gender differences in students' satisfaction with their co-op experiences. Several works observed that satisfaction with the engineering major did not translate directly to pursuing a career in engineering, particularly among women [4]. Moreover, women employed in STEM workplaces left more often than men [144], especially during early careers [243, 132]. The most commonly cited reasons for female attrition included dissatisfaction over pay and working conditions, overt and implicit sexism, gendered expectations, and lack of professionalism [157, 285, 293, 126]. Additionally, gender differences were found in what men and women valued at their jobs [218, 184, 132]. For example, studies found that women who receive more workplace support are more satisfied and stay in engineering longer, indicating that satisfaction can affect retention [120, 10].

Since early workplace experiences can greatly affect subsequent career choices [174], identifying aspects that cause dissatisfaction and providing incentives valued by female employees may encourage more female students to join engineering careers as well as increase retention. Therefore, to identify if male and female students are equally satisfied with their early career work experiences, this section analyzes how they evaluate the various aspects of their co-op work terms (Figures 1.1 and 6.1).

6.5.2 Data and Methods

The analysis is enabled by the Co-op dataset, which contains records of students' experiences as they move from applications to evaluations in the co-op pipeline (refer to Section 3.3 and Figure 1.2). As can be seen in Figure 1.2, at the end of the co-op work terms, students evaluate their workplace experiences. In addition to providing an overall evaluation, students rate various aspects of their work term, including the availability of employer support and the compensation received (the details of the seven specific satisfaction criteria can be found in Section 3.3). Recall that the 2015/2016 Co-op dataset, which is used to measure gender differences in all other stages, contains only overall satisfaction scores and a separate dataset from 2017 is used to analyze detailed satisfaction scores. Following Figure 6.1, we apply statistical analysis to students' evaluations of their work terms to examine any gender differences in satisfaction.

We use the (1) Mann-Whitney test to compare the average overall satisfaction scores of male and female students from the 2015/2016 and 2017 dataset, (2) Mann-Whitney test to compare the average scores provided by male and female students on the seven specific satisfaction criteria in the 2017 dataset (listed in Section 3.3). Again, we use the Mann-Whitney test because of the ordinal nature of the evaluation data and report gender differences in satisfaction for all, junior, and senior students of ENG, COMP, and MECH. Details about the Mann-Whitney test can be found in Section 4.1.

6.5.3 Results

Table 6.19 shows significant differences in co-op students' overall satisfaction (found in the 2015/2016 and the 2017 datasets) as well as the seven detailed satisfaction scores (found in the 2017 dataset); the same notational conventions as in Tables 6.3 and 6.13 are used. We start with overall work term satisfaction scores. In the 2015/2016 dataset, male students appear to be more satisfied in all of ENG and all of COMP. Breaking down by seniority, senior male students in all of ENG and in MECH give higher satisfaction scores, but other groups do not show any significant differences. In the 2017 dataset, overall satisfaction is again higher for all male students, but this trend does not carry over to any subgroups.

Next, we look at gender differences in the seven criteria of student satisfaction. Table 6.19 shows that all criteria except *availability of employer support* either show no difference in satisfaction or a difference in favour of male students. In particular, male students appear to be more satisfied than female students with opportunities to develop their professional network and do work more closely related to their academic program during co-op. Male students, including junior male students, report more opportunities to make meaningful contributions than their female counterparts. On the criterion of receiving appropriate compensation (shown in row five of Table 6.19), senior COMP male students' average score is 0.23 higher (on a scale of 5) than senior COMP female students', which is the highest statistically significant difference of means in this analysis. On the other hand, female co-op students appear to be more satisfied than male students with the availability of employer support. As can be seen in the first row of Table 6.19, COMP students, but not senior female students in isolation, give higher scores on this criterion.

6.5.4 Discussion

Our analysis of students' end-of-term evaluation of their co-op employers suggested gender differences in (perceived) workplace experiences, and in turn, satisfaction. Male students appeared to be more satisfied, especially with opportunities to make meaningful contributions at work, expand their professional network, and work on topics related to what they

Table 6.19: Gender differences in overall work term satisfaction and satisfaction with specific aspects of co-op jobs: 2017 and 2015/2016 dataset

Criteria	All			Junior			Senior		
	ENG	COMP	MECH	ENG	COMP	MECH	ENG	COMP	MECH
Availability of employer support	-	F0.1**	-	-	F0.1*	-	-	-	-
Opportunities to learn/develop new skills	-	-	-	-	-	-	-	-	-
Opportunities to make meaningful contributions	M0.09**	-	-	M0.11**	-	-	-	-	-
Opportunities to expand professional network	M0.06*	M0.08*	M0.13*	-	-	M0.2*	-	M0.21*	-
Appropriate compensation and/or benefits	-	-	-	-	-	-	-	M0.23*	-
Work related to academic program	M0.1***	-	-	M0.12**	-	-	M0.11*	M0.14*	-
Work related to skills developed at university	-	-	-	-	-	-	-	-	-
Overall Satisfaction (2017)	M0.08**	-	-	-	-	-	-	-	-
Overall Satisfaction (2015/2016)	M0.12***	M0.10*	-	-	-	-	M0.18**	-	M0.57**

learned in the classroom. Additionally, senior male students in computing reported greater satisfaction with receiving appropriate compensation than female students. On the other hand, the only difference in favour of female students was in the availability of employer support, observed mainly in junior female students in computing (Table 6.19).

While interviewing co-op students is the only way to uncover the ground truth behind the gender difference in satisfaction scores, we speculate that gender differences in (a) opportunity (or students' perception of the opportunities received), or (b) expectation, could have led to the above observation.

Gender differences in opportunity: Prior work found evidence of men receiving more opportunities than women, especially in terms of compensation, promotion, meaningful work profiles, and workplace support. [17, 251, 285, 293, 126]. Studies suggest that dissatisfaction over pay and working conditions in STEM workplaces could explain the higher rate of attrition for women [157]. In addition, past, largely qualitative, findings suggest that the under-representation of women and the implicit discrimination against them lead to further gender differences in opportunity (or the perception thereof) [285]. However, assuming that "employer support" is related to "mentorship", our finding contradicts past studies, which found that women receive less mentoring than men [17, 232].

Gender differences in expectation: A difference in what men and women expect or value may contribute towards a gender difference in satisfaction. Research suggested that STEM workplaces tend to offer incentives that are valued by men more than women [184]. Moreover, studies suggest that men and women may evaluate jobs and careers on different scales and criteria [184, 308], resulting in the observed differences in satisfaction.

That being said, since satisfaction in early career experiences may affect students' future career choices, attempts should be made to further examine the reasons behind these gender differences in satisfaction and make engineering workplaces more bias-free and inclusive.

6.5.5 Summary and Conclusions

This section analyzes gender differences in students' satisfaction with early engineering careers. Our analysis is enabled by a dataset that contains information about students' co-op experiences, from applications to evaluations. As part of the evaluation stage of the co-op process, students rate their employers on various criteria. Since gender differences in satisfaction can impact young engineers' career trajectories, this section analyzes gender differences in students' satisfaction with their co-op work term experiences. While gender differences in satisfaction with STEM careers have been studied in the past, this is the first work that analyzes gender differences in engineering students' satisfaction with their co-op experiences.

In our analysis, male students appeared to be more satisfied with their co-op experiences, especially with compensation, networking opportunities, and the ability to make meaningful contributions, and female students appeared to be happier with the availability of employer support. In order to attract and retain STEM talent, employers and institutions should ensure that male and female students are satisfied with their work terms. This could be done by ensuring equal opportunity and providing incentives valued by both male and female students. Quarter and mid-semester workplace reviews might be able to help address students' concerns.

While in this thesis we studied employers' evaluations of students and students' satisfaction with their co-op employment separately, it would be interesting to examine the role of gender in the interaction of these scores. As mentioned in the literature review, prior work suggests that women in STEM report negative experiences when working in teams. Thus, investigating the relationship between a student's performance evaluation score on teamwork and their satisfaction with their co-op experience may reveal additional insight. This may require additional knowledge of the nature of the teamwork, for example, the presence of female peers and mentors on the team. Similarly, examining correlations between various evaluation and satisfaction scores during work terms where the student was rated highly but did not reciprocate or vice-versa may shed light upon additional reasons behind their dissatisfaction.

Chapter 7

Conclusions

The gender gap in engineering programs and careers are well known. In this study, we examine gender differences at two stages of the engineering educational pipeline: (1) during undergraduate admissions and (2) during work-integrated education where we focus on co-operative work terms. Our analysis is enabled by large and unique datasets from the engineering faculty of a large North American university. We apply standard statistical and text analysis methods to these datasets to measure the differences between the interests and experiences of male and female students applying to and enrolled in engineering co-op programs. Specifically, we identify differences in the motivations, interests, and high school backgrounds of male and female applicants to engineering co-op programs and measure gender differences in the co-op experiences of enrolled students in terms of their choice, opportunity, perceived competency, and satisfaction.

Most of past work, especially those related to gender differences prior to post-secondary education, are based on small datasets collected through surveys, interviews, and longitudinal studies. In particular, many researchers surveyed male and female secondary school students to understand the reasons behind their choice of major. They found that even among female students who indicated an interest in engineering, many decided against applying to engineering programs due to various reasons. These reasons include the lack of role models, an unsupportive environment, negative stereotypes, discrimination, and misalignment of values. Additionally, analysis of engineering students' career paths post-graduation reveal similar reasons for female attrition. Data analysis of male and female students who have already applied to or are enrolled in engineering programs may either shed light on additional barriers to entry for female students or, at the least, quantify the gender gaps leading to the "leaks" in the pipeline. To the best of our knowledge, our work is the first to provide large-scale data-driven evidence regarding gender differences in the personalities and backgrounds of applicants to undergraduate engineering programs. Additionally, since most research on gender differences in careers has focused on post-graduation employment, our work is the first to identify gender differences in early engineering careers and co-operative education. Besides, our work is also the first to follow competing job candidates through a job search process and identify gender differences in internal employment decisions. Table 7.1 summarizes the main observations of our research. While some of our observations provide data-driven evidence for past findings, others quantify various gender gaps and lead to new actionable insights.

Our findings listed in Table 7.1a suggest that female students still need an "extra push" towards engineering. Among other sources (discussed below), the *fourth row of Table 7.1a* suggests that this push can come in the form of explicit evidence of being good enough at STEM. Therefore, an important strategy to increase the number of female engineering applicants may be to increase female students' confidence in their STEM abilities. Incentivizing student participation in STEM fairs, programs, and out-of-school activities may increase female students' exposure, and in turn, confidence in STEM. Since female high school students appear to have fewer experiences with practical applications of STEM *(first row of Table 7.1a)*, ostensibly because of their collaborative and competitive nature *(fourth row of Table 7.1a)*, providing supplementary STEM resources may help them gain confidence. Moreover, since female students mention contribution towards society as a motivator to join engineering programs *(first row of Table 7.1a)* and undertake various artistic and communal pursuits during high school *(second row of Table 7.1a)*, providing access to STEM extracurricular activities with creative or altruistic goals, may increase the number of female students interested in these activities, and in turn, engineering.

In addition to explicit evidence of being good enough at STEM, our data-driven findings confirm the role of students' surroundings on how it can push them towards engineering programs and careers. Our analysis reveals that female students place greater emphasis on encouragement and guidance from family members, access to role models, and acceptance from peers *(first and fourth rows of Table 7.1a)* and may persist in male-dominated engineering environments if they receive bias-free feedback and special attention while in minority *(third row of Table 7.1b)*. Thus, in order to increase the number of female applicants to engineering programs, institutions may provide career counselling services, accessible role models, and guidance for college admissions, to all high school students. At the same time, efforts should be made towards reducing the pervasive and maybe unconscious pro-male STEM stereotypes held by educators, parents, employers, male peers, and female students themselves. Since the results of this thesis indicate that female students receive more opportunities and better evaluations from their co-op employers *(second and third rows of Table 7.1b)*, they may contribute towards reducing pro-male STEM stereotypes as well as encourage female students to apply to and persist in engineering programs and careers. However, since some of our results also suggest a gender difference in students' evaluation and satisfaction with early career experiences *(third and fourth rows of Table 7.1b)*, they prompt the need to inspect STEM workplaces, especially those that employ young professionals, and highlight the ramifications of gender bias on talent recruitment and retention.

Table 7.1: Summary of observations

(a) Secondary Education

Research Question	Male	Female
Motivation to join engineering	Technical interests (5.1)	Technical interests (5.1)
	Technical applications (5.1)	Family influence, Contribute to society (5.1)
Interests, Extracurricular, Jobs	Mainly technical (5.1)	Variety (art, community, service) (5.1)
Perception of Co-op		Mention co-op more frequently as a reason to join the university (5.2)
	Leverage reputation and size of institution's co-op program (5.2)	Gain new skills, Practical experience, Post-graduation job, Try many career options (5.2)
High school contexts		Greater gender gap in math test scores, with female students performing better (5.3)
		Personal influence and guidance (5.3)
		Variety of subjects (5.3)
		Interest and capability in STEM as opposed to career security as a reason to study engineering (5.3)
		Fewer collaborative and competitive technical activities (5.3)

In addition to suggesting ways of encouraging female students to join and persist in engineering, our findings also suggest that engineering programs and careers be oriented

Table 7.1: Summary of observations, continued

(b) Undergraduate Co-operative Education

<table>
<tr><th>Research Question</th><th>Male</th><th>Female</th></tr>
<tr><td rowspan="2">Choice</td><td colspan="2">Design and analysis (6.1)</td></tr>
<tr><td>Hardware, Manufacturing (6.1)</td><td>User Interface, Project management (6.1)</td></tr>
<tr><td rowspan="3">Opportunity</td><td colspan="2">Equal or in favour of female students (6.2, 6.3)</td></tr>
<tr><td></td><td>Slightly more interviews, shortlists, offers, and placements (6.2, 6.3)</td></tr>
<tr><td></td><td>Riskier ranking strategies (6.2)</td></tr>
<tr><td rowspan="5">Perceived Competency</td><td></td><td>Higher Overall Performance (6.4)</td></tr>
<tr><td>Entrepreneurial, Eager, Efficient (6.4)</td><td>Collaborative, Thorough, Dependable, Adaptable, Communication (6.4)</td></tr>
<tr><td>Recommended to be collaborative and attentive to mistakes (6.4)</td><td>Recommended to increase business knowledge (6.4)</td></tr>
<tr><td colspan="2">Visible majority received technical feedback and recommendation</td></tr>
<tr><td colspan="2">Visible minority was recommended to increase confidence and ask more questions (6.4)</td></tr>
<tr><td rowspan="2">Satisfaction</td><td>Higher Overall Satisfaction (6.5)</td><td></td></tr>
<tr><td>Compensation, Networking opportunities, Meaningful contribution, Relevance to academic program (6.5)</td><td>Employer support (6.5)</td></tr>
</table>

to accommodate the expectations of female students. For instance, our analysis reveals that female students have a breadth of interests *(second and third rows of Table 7.1a)*; female students report participation in a breadth of extracurricular activities, including community service, during high school and list experiencing a variety of career options as a reason to join co-op programs. Moreover, the analysis of co-op applications sent by female students reveals that they prefer a variety of jobs, including jobs that directly affect people, such as user interface design, as well as those that require coordinating with them, for example, project management *(first row of Table 7.1b)*. Thus, a way to attract and retain more female students in engineering would be to present it as a profession that can help others and allow for a broad range of careers and learning opportunities. Highlighting the different types of available jobs, including those preferred by female students, and adding relevant material to engineering curricula may foster a more inclusive image of engineering. This may even lower the observed dissatisfaction of female students regarding how their co-op job was not related to their academic program *(fourth row of Table 7.1b)*.

As shown above, the results presented in this thesis are relevant to students, institutions, and co-op employers. However, they should be interpreted carefully since they are based on datasets from a single North American institution. Even though secondary data analyses of these datasets provide actionable data-driven insight, they also add a limitation to the research. The data analysis answers the “what”, but gives no evidence regarding the “why”. In other words, we can identify interesting patterns and correlations in the data, but it is difficult to explain why they occur without additional analysis or interviewing students and employers to collect additional information. For example, we found that female students in computing send more applications to jobs related to user experience and interface. However, to understand why this happens, students will have to be interviewed and additional data about their co-op experiences will have to be collected. Thus, further analysis and perhaps interviews would be required to confirm any cause-and-effect relationships. Therefore, while this research gives us a starting point for further inquiry into the gender gap in engineering, it raises more questions than it answers.

Apart from the gender differences we have studied, data analysis can be conducted on additional aspects of the STEM educational pipeline. For example, students’ satisfaction with co-op may depend on other factors besides the seven criteria identified by the university. Analyzing interactions between students’ performance evaluation and satisfaction scores may reveal additional reasons behind their dissatisfaction. Moreover, examining undergraduate students’ choices of elective courses, clubs, competitions, and other on-campus activities, might reveal gender differences in interest, and using language models to check and correct gendered language in online evaluation forms may mitigate gender differences in feedback. Furthermore, mining social media channels to study the differences between what male and female students think about STEM while in their secondary and undergraduate education and modelling their mental and social health may provide insights into what affects their STEM aspirations the most. Collecting additional data from alumni could reveal how positive or negative early career experiences impact students' career paths, especially, gender-specific attrition. Finally, expanding the scope of the research to non-STEM programs, specifically those that are female-dominated, and tracing the effect of diverting more female students towards engineering may provide an interesting direction for future work.

References

[1] Lauren B Adamson, Martha A Foster, Martha L Roark, and Donna B Reed. Doing a science project: Gender differences during childhood. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 35(8):845–857, 1998.

[2] Sigal Alon and Thomas A DiPrete. Gender differences in the formation of a field of study choice set. *Sociological science*, 2:50, 2015.

[3] Susan A Ambrose and Laurie Poklop. Do students really learn from experience? *Change: The magazine of higher learning*, 47(1):54–61, 2015.

[4] Catherine T Amelink and Elizabeth G Creamer. Gender differences in elements of the undergraduate experience that influence satisfaction with the engineering major and the intent to pursue engineering as a career. *Journal of Engineering Education*, 99(1):81–92, 2010.

[5] EARL Anderson, NANCY Johnston, LARRY Iles, NORAH Mcrae, NANCY Reed, and JULIE Walchli. Co-operative education and student recruitment, engagement and success: Early findings from a multi-institutional study in british columbia. *Journal of Cooperative Education and Internships*, 46(1):58–76, 2012.

[6] Pauziah Mohd Arsad, Norlida Buniyamin, and JA Manan. Evaluation of academic performance of electrical engineering students based on gender and various entry levels. *Wseas Transactions on Advances in Engineering Education*, 8(2):62–71, 2011.

[7] Cornelia M Ashby. Higher education: Science, technology, engineering, and mathematics trends and the role of federal programs. testimony before the committee on education and the workforce, house of representatives. gao-06-702t. *Government Accountability Office*, 2006.

[8] Elcin Ayaz and Rabia Sarikaya. The effect of engineering design-based science teaching on the perceptions of classroom teacher candidates towards stem disciplines. *International Journal of Progressive Education*, 15(3):13–27, 2019. https://doi.org/10.29329/ijpe.2019.193.2.

[9] Sevgi Aydin-Gunbatar, Aysegul Tarkin-Celikkiran, Elif Selcan Kutucu, and Betul Ekiz-Kiran. The influence of a design-based elective stem course on pre-service chemistry teachers' content knowledge, stem conceptions, and engineering views. *Chemistry Education Research and Practice*, 19(3):954–972, 2018. https://doi.org/10.1039/C8RP00128F.

[10] Mary Ayre, Julie Mills, and Judith Gill. Yes, i do belong: The women who stay in engineering. *Engineering studies*, 5(3):216–232, 2013.

[11] Oleg Badunenko, Nataliya Barasinska, and Dorothea Schäfer. Investments: Women are more cautious than men because they have less financial resources at their disposal. *Weekly Report*, 6(1):1–4, 2010.

[12] Caroline Baillie and Alice Pawley. *Engineering and social justice: In the university and beyond*. Purdue University Press, 2012.

[13] Christopher Ball, Kuo-Ting Huang, Shelia R Cotten, and RV Rikard. Pressurizing the stem pipeline: An expectancy-value theory analysis of youths' stem attitudes. *Journal of Science Education and Technology*, 26(4):372–382, 2017. https://doi.org/10.1007/s10956-017-9685-1.

[14] Alison E Barber and Mark V Roehling. Job postings and the decision to interview: A verbal protocol analysis. *Journal of applied psychology*, 78(5):845, 1993.

[15] David N Beede, Tiffany A Julian, David Langdon, George McKittrick, Beethika Khan, and Mark E Doms. Women in stem: A gender gap to innovation. 2011.

[16] Jeri Mullins Beggs, John H Bantham, and Steven Taylor. Distinguishing the factors influencing college students' choice of major. *College Student Journal*, 42(2):381–395, 2008.

[17] Catherine White Berheide, Lisa Christenson, Rena Linden, and Una Bray. Gender differences in promotion experiences at two elite private liberal arts colleges in the united states. In *Forum on Public Policy Online*, volume 2013. ERIC, 2013.

[18] Marianne Bertrand, Claudia Goldin, and Lawrence F Katz. Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics*, 2(3):228–55, 2010.

[19] Nancy E Betz and Gail Hackett. The relationship of career-related self-efficacy expectations to perceived career options in college women and men. *Journal of counseling psychology*, 28(5):399, 1981.

[20] Sylvia Beyer, Kristina Rynes, and Susan Haller. Deterrents to women taking computer science courses. *IEEE technology and society magazine*, 23(1):21–28, 2004.

[21] Kaninika Bhatnagar, Mary Brake, and Al Bellamy. Choice of technology majors among high school students. *Advancing Women in Leadership*, 25:N_A, 2007.

[22] Benjamin F Blair, Meghan Millea, and Joshua Hammer. The impact of cooperative education on academic performance and compensation of engineering majors. *Journal of Engineering Education*, 93(4):333–338, 2004.

[23] Francine D Blau and Lawrence M Kahn. Gender differences in pay. *Journal of Economic perspectives*, 14(4):75–99, 2000.

[24] Martha M Bleeker and Janis E Jacobs. Achievement in math and science: Do mothers' beliefs matter 12 years later? *Journal of Educational Psychology*, 96(1):97, 2004.

[25] Rae Lesser Blumberg. *Gender bias in textbooks: A hidden obstacle on the road to gender equality in education*. Unesco Paris, 2007.

[26] Werner Bönte and Stefan Krabel. You can't always get what you want: Gender differences in job satisfaction of university graduates. *Applied Economics*, 46(21):2477–2487, 2014.

[27] Alison L Booth, Marco Francesconi, and Jeff Frank. A sticky floors model of promotion, pay, and gender. *European Economic Review*, 47(2):295–322, 2003.

[28] Maura Borrego, Jennifer Karlin, Lisa D McNair, and Kacey Beddoes. Team effectiveness theory from industrial and organizational psychology applied to engineering student project teams: A research review. *Journal of Engineering Education*, 102(4):472–512, 2013.

[29] Laura J Bottomley, Sarah Rajala, and Richard Porter. Women in engineering at north carolina state university: An effort in recruitment, retention, and encouragement. In *FIE'99 Frontiers in Education. 29th Annual Frontiers in Education Conference. Designing the Future of Science and Engineering Education. Conference Proceedings (IEEE Cat. No. 99CH37011*, volume 1, pages 11A5–1. IEEE, 1999.

[30] Kathryn L Boucher, Melissa A Fuesting, Amanda B Diekman, and Mary C Murphy. Can i work with and help others in this field? how communal goals influence interest and participation in stem fields. *Frontiers in psychology*, 8:901, 2017. https://doi.org/10.3389/fpsyg.2017.00901.

[31] Samuel Bowles and Herbert Gintis. The inheritance of inequality. *Journal of economic Perspectives*, 16(3):3–30, 2002. https://doi.org/10.1257/089533002760278686.

[32] Suzanne G Brainard and Linda Carlin. A longitudinal study of undergraduate women in engineering and science. In *Proceedings Frontiers in Education 1997 27th Annual Conference. Teaching and Learning in an Era of Change*, volume 1, pages 134–143. IEEE, 1997.

[33] Glynis M Breakwell and Sue Beardsell. Gender, parental and peer influences upon science attitudes and activities. *Public Understanding of Science*, 2016.

[34] Thomas Breda and Mélina Hillion. Teaching accreditation exams reveal grading biases favor women in male-dominated disciplines in france. *Science*, 353(6298):474–478, 2016.

[35] Alan Brown. What engineering shortage? *IEEE Engineering Management Review*, 37(4):93, 2009. https://doi.org/10.1109/EMR.2009.5384056.

[36] Kingsley Browne. *Biology at work: Rethinking sexual equality*. Rutgers University Press, 2002.

[37] Kingsley R Browne. Evolved sex differences and occupational segregation. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 27(2):143–162, 2006.

[38] Krista Brucker, Nash Whitaker, Zachary S Morgan, Katie Pettit, Erynn Thinnes, Alison M Banta, and Megan M Palmer. Exploring gender bias in nursing evaluations of emergency medicine residents. *Academic Emergency Medicine*, 26(11):1266–1272, 2019.

[39] Tor Busch. Gender differences in self-efficacy and attitudes toward computers. *Journal of Educational Computing Research*, 12(2):147–158, 1995.

[40] Kathleen Buse, Diana Bilimoria, and Sheri Perelli. Why they stay: women persisting in us engineering careers. *Career Development International*, 18(2):139–154, 2013.

[41] James P Byrnes, David C Miller, and William D Schafer. Gender differences in risk taking: a meta-analysis. *Psychological bulletin*, 125(3):367, 1999.

[42] Jill M Bystydzienski, Margaret Eisenhart, and Monica Bruning. High school is not too late: Developing girls' interest and engagement in engineering careers. *The Career Development Quarterly*, 63(1):88–95, 2015.

[43] Robert Cabral, Marianne A Ferber, and Carole A Green. Men and women in fiduciary institutions: A study of sex differences in career development. *The Review of Economics and Statistics*, pages 573–580, 1981.

[44] CAFCE. Canadian association for co-operative education. http://www.cafce.ca/accreditation.html. Accessed: 20th August, 2017.

[45] Scott E Carrell, Marianne E Page, and James E West. Sex and science: How professor gender perpetuates the gender gap. *The Quarterly Journal of Economics*, 125(3):1101–1144, 2010.

[46] Sophia Catsambis. Gender, race, ethnicity, and science education in the middle grades. *Journal of Research in Science Teaching*, 32(3):243–257, 1995.

[47] Paola Cecchi-Dimeglio. How gender bias corrupts performance reviews, and what to do about it. *Harvard Business Review*, 12, 2017.

[48] Stephen J Ceci and Wendy M Williams. Understanding current causes of women's underrepresentation in science. *Proceedings of the National Academy of Sciences*, 108(8):3157–3162, 2011.

[49] Stephen J Ceci and Wendy M Williams. Women have substantial advantage in stem faculty hiring, except when competing against more-accomplished men. *Frontiers in psychology*, 6:1532, 2015.

[50] Sharon K Chambers and Thomas Andre. Gender, prior knowledge, interest, and experience in electricity and conceptual change text manipulations in learning about direct current. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 34(2):107–123, 1997.

[51] Susan Chanderbhan-Forde, Rebekah S Heppner, and Kathryn M Borman. " the doors are open" but they don't come in: Cultural capital and the pathway to engineering degrees for women. *Journal of Women and Minorities in Science and Engineering*, 18(2), 2012.

[52] Mitchell J Chang, Jessica Sharkness, Sylvia Hurtado, and Christopher B Newman. What matters in college for retaining aspiring scientists and engineers from underrepresented racial groups. *Journal of Research in Science Teaching*, 51(5):555–580, 2014.

[53] John C Chen, Samuel Owusu-Ofori, Devdas Pai, Eugenia Toca-McDowell, Shih-Liang Wang, and Cynthia K Waters. A study of female academic performance in mechanical engineering. In *Technology-Based Re-Engineering Engineering Education Proceedings of Frontiers in Education FIE'96 26th Annual Conference*, volume 2, pages 779–782. IEEE, 1996.

[54] Sapna Cheryan, Andrew N Meltzoff, and Saenam Kim. Classrooms matter: The design of virtual classrooms influences gender disparities in computer science classes. *Computers & Education*, 57(2):1825–1835, 2011.

[55] Sapna Cheryan and Victoria C Plaut. Explaining underrepresentation: A theory of precluded interest. *Sex roles*, 63(7-8):475–488, 2010.

[56] Sapna Cheryan, Victoria C Plaut, Paul G Davies, and Claude M Steele. Ambient belonging: How stereotypical cues impact gender participation in computer science. *Journal of personality and social psychology*, 97(6):1045, 2009.

[57] Sapna Cheryan, John Oliver Siy, Marissa Vichayapai, Benjamin J Drury, and Saenam Kim. Do female and male role models who embody stem stereotypes hinder women's anticipated success in stem? *Social Psychological and Personality Science*, 2(6):656–664, 2011.

[58] Naomi C Chesler and Mark A Chesler. Gender-informed mentoring strategies for women engineering scholars: On establishing a caring community. *Journal of engineering education*, 91(1):49–55, 2002.

[59] Shivangi Chopra. Job description mining to understand undergraduate co-operative placements. Master's thesis, University of Waterloo, 2017.

[60] Shivangi Chopra, Hannah Gautreau, Abeer Khan, Melicaalsadat Mirsafian, and Lukasz Golab. Gender differences in undergraduate engineering applicants: A text

mining approach. In *Proceedings of the International Conference on Educational Data Mining (EDM)*, pages 44–54. ERIC, 2018.

[61] Shivangi Chopra and Lukasz Golab. Job description mining to understand work-integrated learning. In *Proceedings of the 11th International Conference on Educational Data Mining, EDM 2018, Buffalo, NY, USA, July 15-18, 2018*, pages 32–43, 2018.

[62] Shivangi Chopra and Lukasz Golab. Forthcoming: Analyzing ranking strategies to characterize competition in the co-operative education job market. *International Journal of Work-Integrated Learning*, 2022.

[63] Shivangi Chopra and Lukasz Golab. Gender differences in early career performance reviews: a text mining study. In *EDBT/ICDT Workshops*, 2022.

[64] Shivangi Chopra and Lukasz Golab. Undergraduate engineering applicants' perceptions of cooperative education: A text mining approach. *International Journal of Work-Integrated Learning*, 23(1):95, 2022.

[65] Shivangi Chopra, Abeer Khan, Melicaalsadat Mirsafian, and Lukasz Golab. Gender differences in work-integrated learning assessments. In *Proceedings of the International Conference on Educational Data Mining (EDM)*, pages 524–527. ERIC, 2019.

[66] Shivangi Chopra, Abeer Khan, Melicaalsadat Mirsafian, and Lukasz Golab. Gender differences in work-integrated learning experiences of stem students: From applications to evaluations. *International Journal of Work-Integrated Learning*, 21(3):253–274, 2020.

[67] Shivangi Chopra, Melicaalsadat Mirsafian, Abeer Khan, and Lukasz Golab. Gender differences in science and engineering: A data mining approach. In *EDBT/ICDT Workshops*, 2019.

[68] Vasilia Christidou. Greek students' science-related interests and experiences: Gender differences and correlations. *International Journal of Science Education*, 28(10):1181–1199, 2006.

[69] Jacob Clark Blickenstaff*. Women and science careers: leaky pipeline or gender filter? *Gender and education*, 17(4):369–386, 2005.

[70] Deborah A Cobb-Clark. Getting ahead: the determinants of and payoffs to internal promotion for young us men and women. In *Worker wellbeing in a changing labor market*, pages 339–372. Emerald Group Publishing Limited, 2001.

[71] Richard K Coll, Karsten Zegwaard, and Dave Hodges. Science and technology stakeholders' ranking of graduate competencies part 1: Employers perspective. *International Journal of Work-Integrated Learning*, 3(2):19, 2002.

[72] Richard K Coll and Karsten E Zegwaard. Perceptions of desirable graduate competencies for science and technology new graduates. *Research in Science & Technological Education*, 24(1):29–58, 2006.

[73] James P Concannon and Lloyd H Barrow. Men's and women's intentions to persist in undergraduate engineering degree programs. *Journal of Science Education and Technology*, 19(2):133–145, 2010.

[74] Lesley Cooper, Janice Orrell, and Margaret Bowden. *Work integrated learning: A guide to effective practice*. Routledge, 2010.

[75] Robyn Cooper and Carol Heaverlo. Problem solving and creativity and design: What influence do they have on girls' interest in stem subject areas?. *American Journal of Engineering Education*, 4(1):27–38, 2013. https://doi.org/10.19030/ajee.v4i1.7856.

[76] Shelley J Correll. Gender and the career choice process: The role of biased self-assessments. *American journal of Sociology*, 106(6):1691–1730, 2001.

[77] Shelley J Correll, Katherine R Weisshaar, Alison T Wynn, and JoAnne Delfino Wehner. Inside the black box of organizational life: The gendered language of performance assessment. *American Sociological Review*, 85(6):1022–1050, 2020.

[78] Jennifer Crocker, Andrew Karpinski, Diane M Quinn, and Sara K Chase. When grades determine self-worth: consequences of contingent self-worth for male and female engineering and psychology majors. *Journal of personality and social psychology*, 85(3):507, 2003.

[79] W Bruce Croft, Donald Metzler, and Trevor Strohman. *Search engines: Information retrieval in practice*. Addison-Wesley Reading, 2010.

[80] Robert Crosnoe, Catherine Riegle-Crumb, Sam Field, Kenneth Frank, and Chandra Muller. Peer group contexts of girls' and boys' academic experiences. *Child Development*, 79(1):139–155, 2008.

[81] Jan Cuny and William Aspray. Recruitment and retention of women graduate students in computer science and engineering: results of a workshop organized

by the computing research association. *SIGCSE Bulletin*, 34:168–174, 06 2002. https://doi.org/10.1145/543812.543852.

[82] R Darolia, JB Main, C Koedel, JF Ndashimye, and J Yan. The diversity of college engineering degrees: The role of geography and the concentration of engineering degree production. In *ASEE Annual Conference and Exhibition*, 2018.

[83] Rajeev Darolia, Cory Koedel, Joyce B Main, J Felix Ndashimye, and Junpeng Yan. High school course access and postsecondary stem enrollment and attainment. *Educational Evaluation and Policy Analysis*, 42(1):22–45, 2020.

[84] Nilanjana Dasgupta, Melissa McManus Scircle, and Matthew Hunsinger. Female peers in small work groups enhance women's motivation, verbal participation, and career aspirations in engineering. *Proceedings of the National Academy of Sciences*, 112(16):4988–4993, 2015. https://doi.org/10.1073/pnas.1422822112.

[85] Nilanjana Dasgupta and Jane G Stout. Girls and women in science, technology, engineering, and mathematics: Steming the tide and broadening participation in stem careers. *Policy Insights from the Behavioral and Brain Sciences*, 1(1):21–29, 2014. https://doi.org/10.1177/2372732214549471.

[86] Eric D Deemer, Dustin B Thoman, Justin P Chase, and Jessi L Smith. Feeling the threat: Stereotype threat as a contextual barrier to women's science career choice intentions. *Journal of Career Development*, 41(2):141–158, 2014.

[87] Amanda B Diekman, Elizabeth R Brown, Amanda M Johnston, and Emily K Clark. Seeking congruity between goals and roles: A new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychological Science*, 21(8):1051–1057, 2010.

[88] Amanda B Diekman, Emily K Clark, Amanda M Johnston, Elizabeth R Brown, and Mia Steinberg. Malleability in communal goals and beliefs influences attraction to stem careers: Evidence for a goal congruity perspective. *Journal of personality and social psychology*, 101(5):902, 2011.

[89] Chris Ding and Xiaofeng He. K-means clustering via principal component analysis. In *Proceedings of the twenty-first international conference on Machine learning*, page 29. ACM, 2004.

[90] Alessandra Jayne Dinin. An exploration in theory of the storied experiences of women earning engineering bachelor's degrees at a southern, research, predominately white institution. 2017.

[91] Elena Doldor, Madeleine Wyatt, and Jo Silvester. Statesmen or cheerleaders? using topic modeling to examine gendered messages in narrative developmental feedback for leaders. *The Leadership Quarterly*, 30(5):101308, 2019.

[92] David W Drewery, Lauren A Cormier, T Judene Prettti, and Dana Church. Improving unmatched co-op students' emotional wellbeing: Test of two brief interventions. *International Journal of Work-Integrated Learning*, 20(1):43–53, 2019.

[93] Sheelagh Drudy and Máire Úi Chatháin. Gender effects in classroom interaction: Data collection, self-analysis and reflection. *Evaluation & Research in Education*, 16(1):34–50, 2002.

[94] Benjamin J Drury, John Oliver Siy, and Sapna Cheryan. When do female role models benefit women? the importance of differentiating recruitment from retention in stem. *Psychological Inquiry*, 22(4):265–269, 2011.

[95] Maureen Drysdale, Johni Goyder, Amandai Nosko, M Easton, Kristyn Frank, and Patriciai Rowe. The role of co-op in the transition from high school to post-secondary education. *Journal of Cooperative Education and Internships*, 41(1):48–55, 2007.

[96] Kuheli Dutt, Danielle L Pfaff, Ariel F Bernstein, Joseph S Dillard, and Caryn J Block. Gender differences in recommendation letters for postdoctoral fellowships in geoscience. *Nature Geoscience*, 9(11):805, 2016.

[97] Carol S Dweck. The development of ability conceptions. In *Development of achievement motivation*, pages 57–88. Elsevier, 2002.

[98] Chris Eames and Richard K Coll. Cooperative education: Integrating classroom and workplace learning. In *Learning through practice*, pages 180–196. Springer, 2010.

[99] Jacquelynne Eccles. Where are all the women? gender differences in participation in physical science and engineering. In *Why aren't more women in science?: Top researchers debate the evidence*, pages 199–210. American Psychological Association, 2007.

[100] Jacquelynne Eccles. Who am i and what am i going to do with my life? personal and collective identities as motivators of action. *Educational Psychologist*, 44(2):78–89, 2009.

[101] Jacquelynne S Eccles, Bonnie Barber, and Debra Jozefowicz. Linking gender to educational, occupational, and recreational choices: Applying the eccles et al. model of achievement-related choices. 1999.

[102] Jacquelynne S Eccles and Ming-Te Wang. What motivates females and males to pursue careers in mathematics and science? *International Journal of Behavioral Development*, 40(2):100–106, 2016.

[103] Kimberly A Eddleston, David C Baldridge, and John F Veiga. Toward modeling the predictors of managerial career success: does gender matter? *Journal of Managerial Psychology*, 19(4):360–385, 2004.

[104] Donna Eder and Stephen Parker. The cultural production and reproduction of gender: The effect of extracurricular activities on peer-group culture. *Sociology of Education*, pages 200–213, 1987.

[105] Mark E Engberg and Gregory C Wolniak. Examining the effects of high school contexts on postsecondary enrollment. *Research in Higher Education*, 51(2):132–153, 2010.

[106] Nathan Ensmenger. Making programming masculine. *Gender codes: Why women are leaving computing*, pages 115–141, 2010.

[107] Nathan L Ensmenger. *The computer boys take over: Computers, programmers, and the politics of technical expertise*. Mit Press, 2012.

[108] Lorelle Espinosa. Pipelines and pathways: Women of color in undergraduate stem majors and the college experiences that contribute to persistence. *Harvard Educational Review*, 81(2):209–241, 2011.

[109] Lorelle L Espinosa. The academic self-concept of african american and latina (o) men and women in stem majors. *Journal of Women and Minorities in Science and Engineering*, 14(2), 2008.

[110] Lynn Farrell and Louise McHugh. Examining gender-stem bias among stem and non-stem students using the implicit relational assessment procedure (irap). *Journal of Contextual Behavioral Science*, 6(1):80–90, 2017.

[111] Mary Feeney and Margarita Bernal. Women in stem networks: who seeks advice and support from women scientists? *Scientometrics*, 85(3):767–790, 2010.

[112] Helga Fehr-Duda, Manuele De Gennaro, and Renate Schubert. Gender, financial risk, and probability weights. *Theory and decision*, 60(2-3):283–313, 2006.

[113] Richard M Felder, Gary N Felder, Meredith Mauney, Charles E Hamrin Jr, and E Jacquelin Dietz. A longitudinal study of engineering student performance and retention. iii. gender differences in student performance and attitudes. *Journal of Engineering Education*, 84(2):151–163, 1995.

[114] Mingyu Feng, Jeremy Roschelle, Craig Mason, and Ruchi Bhanot. Investigating gender differences on homework in middle school mathematics. In *Proc. of the Int. Conf. on Educational Data Mining (EDM)*, pages 364–369, 2016.

[115] Roberto M Fernandez and Santiago Campero. Gender sorting and the glass ceiling in high-tech firms. *ILR Review*, 70(1):73–104, 2017.

[116] Sonia Ferns, Matthew Campbell, and Karsten E Zegwaard. Work integrated learning. 2014.

[117] Sonia Ferns and Keri Moore. Assessing student outcomes in fieldwork placements: An overview of current practice. *Asia-Pacific Journal of Cooperative Education*, 13(4):207–224, 2012.

[118] Kimberley Ferriman, David Lubinski, and Camilla P Benbow. Work preferences, life values, and personal views of top math/science graduate students and the profoundly gifted: Developmental changes and gender differences during emerging adulthood and parenthood. *Journal of personality and social psychology*, 97(3):517, 2009.

[119] Priorities for Science in the 21st Century for the Asia-Pacific Region. Women and gender, science, engineering and technology. http://www.unesco.org/science/wcs/meetings/apa_sydney_e_98.htm, 1998. Accessed: 20th March, 2019.

[120] Nadya A Fouad, Romila Singh, Kevin Cappaert, Wen-hsin Chang, and Min Wan. Comparison of women engineers who persist in or depart from engineering. *Journal of Vocational Behavior*, 92:79–93, 2016.

[121] National Science Foundation. Women, minorities, and persons with disabilities in science and engineering, 1999.

[122] National Science Foundation. Women, minorities, and persons with disabilities in science and engineering, 2011.

[123] Jennifer A Fredricks and Jacquelynne S Eccles. Extracurricular involvement and adolescent adjustment: Impact of duration, number of activities, and breadth of participation. *Applied developmental science*, 10(3):132–146, 2006. https://doi.org/10.1207/s1532480xads1003_3.

[124] Alexandra M Freund, David Weiss, and Bettina S Wiese. Graduating from high school: The role of gender-related attitudes, self-concept and goal clarity in a major transition in late adolescence. *European Journal of Developmental Psychology*, 10(5):580–596, 2013.

[125] Melissa A Fuesting, Amanda B Diekman, and Lynette Hudiburgh. From classroom to career: The unique role of communal processes in predicting interest in stem careers. *Social Psychology of Education*, 20(4):875–896, 2017. https://doi.org/10.1007/s11218-017-9398-6.

[126] Susan K Gardner and Amy Blackstone. Putting in your time: Faculty experiences in the process of promotion to professor. *Innovative Higher Education*, 38(5):411–425, 2013.

[127] Jack Gault, John Redington, and Tammy Schlager. Undergraduate business internships and career success: are they related? *Journal of marketing education*, 22(1):45–53, 2000.

[128] Marcia Gentry, Robert K Gable, and Mary G Rizza. Students' perceptions of classroom activities: are there grade-level and gender differences? *Journal of Educational Psychology*, 94(3):539, 2002.

[129] Aneta George and Liam Peyton. Work-term experience and graduate attribute assessment in software engineering co-op program at university of ottawa. *Proceedings of the Canadian Engineering Education Association (CEEA)*, 2016.

[130] Barry Gerhart. Gender differences in current and starting salaries: The role of performance, college major, and job title. *ILR Review*, 43(4):418–433, 1990.

[131] Andrew L Gillen, Timothy Kinoshita, David Knight, Jacob Grohs, Cheryl Carrico, Holly M Matusovich, Isabel Bradburn, and Lin Tan. Gatekeepers to broadening participation in engineering investigating variation across high schools comparing who could go versus who does go into engineering. In *2017 IEEE Frontiers in Education Conference (FIE)*, pages 1–5. IEEE, 2017.

[132] Jennifer L Glass, Sharon Sassler, Yael Levitte, and Katherine M Michelmore. What's so special about stem? a comparison of women's retention in stem and professional occupations. *Social forces*, 92(2):723–756, 2013.

[133] Stephen Gorard and Beng Huat See. The impact of socio-economic status on participation and attainment in science. *Studies in Science Education*, 45(1):93–129, 2009. https://doi.org/10.1080/03057260802681821.

[134] Eric D Gould and Avi Simhon. Does quality time produce quality children? evidence on the intergenerational transmission of human capital using parental deaths. 2011. https://doi.org/10.2139/ssrn.1765652.

[135] Robert E Gould. Men, money and masculinity. In *Money and mind*, pages 61–66. Springer, 1991.

[136] Kimberly A Goyette and Ann L Mullen. Who studies the arts and sciences? social background and the choice and consequences of undergraduate field of study. *The Journal of Higher Education*, 77(3):497–538, 2006.

[137] Amanda L Griffith. Persistence of women and minorities in stem field majors: Is it the school that matters? *Economics of Education Review*, 29(6):911–922, 2010.

[138] Larue Guyer and Linda Fidell. Publications of men and women psychologists: Do women publish less? *American Psychologist*, 28(2):157, 1973.

[139] Meara M Habashi, William G Graziano, Demetra Evangelou, and Ida Ngambeki. Teacher influences on child interest in stem careers. In *Proceedings of the Research in Engineering Education Symposium*, pages 20–23. Citeseer, 2009.

[140] Mahmoud Haddara and Heather Skanes. A reflection on cooperative education: From experience to experiential learning. *International Journal of Work-Integrated Learning*, 8(1):67, 2007.

[141] Frances Montgomery Haemmerlie and Robert L Montgomery. Gender differences in the academic performance and retention of undergraduate engineering majors. *College Student Journal*, 46(1):40–46, 2012.

[142] Catherine Hakim. Women, careers, and work-life preferences. *British Journal of Guidance & Counselling*, 34(3):279–294, 2006.

[143] Cathy Hall, Jeremy Dickerson, David Batts, Paul Kauffmann, and Michael Bosse. Are we missing opportunities to encourage interest in stem fields?. *Journal of Technology Education*, 23(1):32–46, 2011. https://doi.org/10.21061/jte.v23i1.a.4.

[144] Darcy Hango. Gender differences in science, technology, engineering, mathematics, and computer science (STEM) programs at university. *Insights on Canadian Society*, 12 2013.

[145] Kent D Harber. Feedback to minorities: Evidence of a positive bias. *Journal of personality and social psychology*, 74(3):622, 1998.

[146] Russel S Hathaway, Sally Sharp, and Cinda-Sue Davis. Programmatic efforts affect retention of women in science and engineering. *Journal of Women and Minorities in Science and Engineering*, 7(2), 2001.

[147] Jay Hays and Mike Clements. Supervision in work experience for learning programs. In *Proceedings of the 17th World Conference on Cooperative and Work-Integrated Education (WACE)*, 2011.

[148] Madeline E Heilman. Gender stereotypes and workplace bias. *Research in organizational Behavior*, 32:113–135, 2012.

[149] Cher C Hendricks, Meltem Alemdar, and Tamra Williams Ogletree. The impact of participation in vex robotics competition on middle and high school students' interest in pursuing stem studies and stem-related careers. In *2012 ASEE Annual Conference & Exposition*, pages 25–1312, 2012. https://doi.org/10.18260/1-2--22069.

[150] Joni Hersch and W Kip Viscusi. Gender differences in promotions and wages. *Industrial Relations: A Journal of Economy and Society*, 35(4):461–472, 1996.

[151] Marion Hersh. The changing position of women in engineering worldwide. *IEEE Transactions on Engineering Management*, 47(3):345–359, 2000.

[152] Catherine Hill, Christianne Corbett, and Andresse St Rose. *Why so few? Women in science, technology, engineering, and mathematics.* ERIC, 2010.

[153] Antentor O Hinton Jr, Christina M Termini, Elsie C Spencer, Florentine UN Rutaganira, Daphney Chery, ReAna Roby, Zer Vue, Angela D Pack, Lillian J Brady, and Edgar Garza-Lopez. Patching the leaks: revitalizing and reimagining the stem pipeline. *Cell*, 183(3):568–575, 2020.

[154] Dave Hodges and Noel Burchell. Business graduate competencies: Employers' views on importance and performance. *International Journal of Work-Integrated Learning*, 4(2):16, 2003.

[155] Randy Hodson. Gender differences in job satisfaction: why aren't women more dissatisfied? *The Sociological Quarterly*, 30(3):385–399, 1989.

[156] Shouping Hu and Gregory C Wolniak. College student engagement and early career earnings: Differences by gender, race/ethnicity, and academic preparation. *The Review of Higher Education*, 36(2):211–233, 2013.

[157] Jennifer Hunt. Why do women leave science and engineering? *ILR Review*, 69(1):199–226, 2016.

[158] Anne-Barrie Hunter and Elaine Seymour. *Talking about Leaving Revisited: Persistence, Relocation, and Loss in Undergraduate STEM Education.* Springer International Publishing, 2019.

[159] Sadiq Hussain, Jiten Hazarika, and Pranjal Buragohain. Educational data mining on performance of under graduate students of dibrugarh university using r. *International Journal of Computer Applications*, 114(11):10–16, 2015.

[160] Janet S Hyde, Sara M Lindberg, Marcia C Linn, Amy B Ellis, and Caroline C Williams. Gender similarities characterize math performance. *Science*, 321(5888):494–495, 2008.

[161] Janet Shibley Hyde. Meta-analysis and the psychology of gender differences. *Signs: Journal of Women in Culture and Society*, 16(1):55–73, 1990.

[162] Michael Inzlicht and Talia Ben-Zeev. A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science*, 11(5):365–371, 2000.

[163] Denise Jackson. Exploring the challenges experienced by international students during work-integrated learning in australia. *Asia Pacific Journal of Education*, 37(3):344–359, 2017.

[164] Denise A Jackson. Using work-integrated learning to enhance career planning among business undergraduates. *Australian Journal of Career Development*, 26(3):153–164, 2017.

[165] Janis E Jacobs and Sandra D Simpkins. Mapping leaks in the math, science, and technology pipeline, 2005.

[166] Yuheng Helen Jiang and Lukasz Golab. On competition for undergraduate co-op placements: A graph mining approach. In *EDM*, pages 394–399, 2016.

[167] Yuheng Helen Jiang, Sally Wai Yin Lee, and Lukasz Golab. Analyzing student and employer satisfaction with cooperative education through multiple data sources. *Asia-Pacific Journal of Cooperative Education*, 16(4):225–240, 2015.

[168] David R Jones and Gerald H Makepeace. Equal worth, equal opportunities: pay and promotion in an internal labour market. *The Economic Journal*, 106(435):401–409, 1996.

[169] M Gail Jones, Ann Howe, and Melissa J Rua. Gender differences in students' experiences, interests, and attitudes toward science and scientists. *Science education*, 84(2):180–192, 2000.

[170] Lois Joy. Salaries of recent male and female college graduates: Educational and labor market effects. *ILR Review*, 56(4):606–621, 2003.

[171] Deborah Kaminski and Cheryl Geisler. Survival analysis of faculty retention in science and engineering by gender. *science*, 335(6070):864–866, 2012.

[172] Hosun Kang, Angela Calabrese Barton, Edna Tan, Sandra D Simpkins, Hyang-yon Rhee, and Chandler Turner. How do middle school girls of color develop stem identities? middle school girls' participation in science activities and identification with stem careers. *Science Education*, 103(2):418–439, 2019. https://doi.org/10.1002/sce.21492.

[173] Rosabeth Moss Kanter. Some effects of proportions on group life. In *The gender gap in psychotherapy*, pages 53–78. Springer, 1977.

[174] Antti Kauhanen and Sami Napari. Gender differences in careers. *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, (117/118):61–88, 2015.

[175] Kristen Keith and Abagail McWilliams. The returns to mobility and job search by gender. *ILR Review*, 52(3):460–477, 1999.

[176] Latonia Haney Keith. Visibility invisibility: Feedback bias in the legal profession. *J. Gender Race & Just.*, 23:315, 2020.

[177] Caitlin Kelleher, Randy Pausch, Randy Pausch, and Sara Kiesler. Storytelling alice motivates middle school girls to learn computer programming. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1455–1464. ACM, 2007.

[178] Grace Kena, Lauren Musu-Gillette, Jennifer Robinson, Xiaolei Wang, Amy Rathbun, Jijun Zhang, Sidney Wilkinson-Flicker, Amy Barmer, and Erin Dunlop Velez Velez. The condition of education 2015. nces 2015-144. *National Center for Education Statistics*, 2015.

[179] Oguzhan Kirdök and Esranur Harman. High school students' career decision-making difficulties according to locus of control. *Universal Journal of Educational Research*, 6(2):242–248, 2018. https://doi.org/10.13189/ujer.2018.060205.

[180] David B Knight, Jacob R Grohs, Isabel S Bradburn, Timothy J Kinoshita, Stacey Vaziri, Holly M. Matusovich, and Cheryl Carrico. Illuminating inequality in access: Variation in enrollment in undergraduate engineering programs across virginia's high schools. *Journal of Engineering Education*, 109(4):665–684, 2020.

[181] David B Knight, JR Grohs, HM Matusovich, IS Bradburn, C Carrico, KJ Chew, SL Vaziri, TJ Kinoshita, and L Tan. Gatekeepers to broadening participation in engineering: Variation in postsecondary engineering-going across virginia's high schools. In *ASEE Annual Conference proceedings*, 2019.

[182] Silvia Knobloch-Westerwick, Carroll J Glynn, and Michael Huge. The matilda effect in science communication: an experiment on gender bias in publication quality perceptions and collaboration interest. *Science Communication*, 35(5):603–625, 2013.

[183] David A Kolb. *Experiential learning: Experience as the source of learning and development*. FT press, 2014.

[184] Alison M Konrad, J Edgar Ritchie Jr, Pamela Lieb, and Elizabeth Corrigall. Sex differences and similarities in job attribute preferences: a meta-analysis. *Psychological bulletin*, 126(4):593, 2000.

[185] Zlatko J. Kovacic. Early prediction of student success: Mining students enrollment data. In *Proc. of Informing Science & IT Education Conference (InSITE)*, 2010.

[186] Norris Krueger Jr and Peter R Dickson. How believing in ourselves increases risk taking: Perceived self-efficacy and opportunity recognition. *Decision sciences*, 25(3):385–400, 1994.

[187] Astrid Kunze. Gender differences in entry wages and early career wages. *Annales d'Economie et de Statistique*, pages 245–265, 2003.

[188] Campbell Leaper, Timea Farkas, and Christia Spears Brown. Adolescent girls' experiences and gender-related beliefs in relation to their motivation in math/science and english. *Journal of youth and adolescence*, 41(3):268–282, 2012.

[189] James Daniel Lee. More than ability: Gender and personal relationships influence science and technology involvement. *Sociology of Education*, pages 349–373, 2002.

[190] Matthew Lee and Laura Huang. Gender bias, social impact framing, and evaluation of entrepreneurial ventures. *Organization Science*, 29(1):1–16, 2018.

[191] Walter C Lee. Pipelines, pathways, and ecosystems: An argument for participation paradigms, 2019.

[192] Karen Leppel, Mary L Williams, and Charles Waldauer. The impact of parental occupation and socioeconomic status on choice of college major. *Journal of Family and Economic issues*, 22(4):373–394, 2001.

[193] Larry L Leslie, Gregory T McClure, and Ronald L Oaxaca. Women and minorities in science and engineering: A life sequence analysis. *The journal of higher education*, 69(3):239–276, 1998.

[194] Sarah-Jane Leslie, Andrei Cimpian, Meredith Meyer, and Edward Freeland. Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, 347(6219):262–265, 2015.

[195] Harry Leung. The impact of information quality of job descriptions on an applicant's decision to pursue a job. Master's thesis, 2007.

[196] Lam B Li, M. Cooperative learning: The active classroom. *The Hong Kong Institute of Education*, 2005.

[197] Bo Lin, Bissan Ghaddar, and Ada Hurst. Text mining undergraduate engineering programs' applications: the role of gender, nationality, and socio-economic status. *arXiv preprint arXiv:2107.14034*, 2021.

[198] Sara M Lindberg, Janet Shibley Hyde, Jennifer L Petersen, and Marcia C Linn. New trends in gender and mathematics performance: a meta-analysis. *Psychological bulletin*, 136(6):1123, 2010.

[199] Pamela J Loprest. Gender differences in wage growth and job mobility. *The American Economic Review*, 82(2):526–532, 1992.

[200] Judith Lorber and Susan A Farrell. *The social construction of gender*. Sage Newbury Park, CA, 1991.

[201] Sarah T Lubienski, Joseph P Robinson, Corinna C Crane, and Colleen M Ganley. Girls' and boys' mathematics achievement, affect, and experiences: Findings from ecls-k. *Journal for Research in Mathematics Education*, 44(4):634–645, 2013.

[202] David Lubinski and Camilla Persson Benbow. Study of mathematically precocious youth after 35 years: Uncovering antecedents for the development of math-science expertise. *Perspectives on psychological science*, 1(4):316–345, 2006.

[203] Valerie C Lundy-Wagner, Cindy P Veenstra, Marisa K Orr, Nichole M Ramirez, Matthew W Ohland, and Russell A Long. Gaining access or losing ground? socioeconomically disadvantaged students in undergraduate engineering, 1994—2003. *The Journal of Higher Education*, 85(3):339–369, 2014.

[204] Darrell Anthony Luzzo and Barbara E Ward. The relative contributions of self-efficacy and locus of control to the prediction of vocational congruence. *Journal of Career Development*, 21(4):307–317, 1995. https://doi.org/10.1177/089484539502100404.

[205] David MacPhee, Samantha Farro, and Silvia Sara Canetto. Academic self-efficacy and performance of underrepresented stem majors: Gender, ethnic, and social class patterns. *Analyses of Social Issues and Public Policy*, 13(1):347–369, 2013.

[206] Joyce B Main, Amanda L Griffith, Xinrui Xu, and Alexandra M Dukes. Choosing an engineering major: A conceptual model of student pathways into engineering. *Journal of Engineering Education*, 111(1):40–64, 2022.

[207] Cathrine Maiorca, Thomas Roberts, Christa Jackson, Sarah Bush, Ashley Delaney, Margaret J Mohr-Schroeder, and Soledad Yao Soledad. Informal learning environments and impact on interest in stem careers. *International Journal of Science and Mathematics Education*, 19(1):45–64, 2021. https://doi.org/10.1007/s10763-019-10038-9.

[208] Adam V Maltese and Robert H Tai. Eyeballs in the fridge: Sources of early interest in science. *International Journal of Science Education*, 32(5):669–685, 2010.

[209] Adam V Maltese and Robert H Tai. Pipeline persistence: Examining the association of educational experiences with earned degrees in stem among us students. *Science education*, 95(5):877–907, 2011.

[210] Jon K Maner, Matthew T Gailliot, David A Butz, and B Michelle Peruche. Power, risk, and the status quo: Does power promote riskier or more conservative decision making? *Personality and Social Psychology Bulletin*, 33(4):451–462, 2007. https://doi.org/10.1177/0146167206297405.

[211] Alan Manning and Joanna Swaffield. The gender gap in early-career wage growth. *The Economic Journal*, 118(530):983–1024, 2008.

[212] Katherine Cumings Mansfield, Anjalé D Welton, and Margaret Grogan. "truth or consequences": a feminist critical policy analysis of the stem crisis. *International Journal of Qualitative Studies in Education*, 27(9):1155–1182, 2014.

[213] Jane Margolis and Allen Fisher. Unlocking the clubhouse: Women in computing cambridge, 2002.

[214] Rose M Marra, Kelly A Rodgers, Demei Shen, and Barbara Bogue. Women engineering students and self-efficacy: A multi-year, multi-institution study of women engineering student self-efficacy. *Journal of engineering education*, 98(1):27–38, 2009. https://doi.org/10.1002/j.2168-9830.2009.tb01003.x.

[215] Douglas J Marshall. Why a co-op program—a student's point of view. *Engineering Issues: Journal of Professional Activities*, 101(2):243–246, 1975.

[216] Carol Lynn Martin. New directions for investigating children s gender knowledge. *Developmental Review*, 13(2):184–204, 1993.

[217] Gary R Martin. A case study of student recruitment and educational implications for colleges considering mandatory co-op. *Journal of Cooperative Education*, 32(3):96–101, 1997.

[218] E Sharon Mason. Gender differences in job satisfaction. *The Journal of Social Psychology*, 135(2):143–151, 1995.

[219] Mary Ann Mason and Marc Goulden. Marriage and baby blues: Redefining gender equity in the academy. *The Annals of the American Academy of Political and Social Science*, 596(1):86–103, 2004.

[220] Holly Matusovich, Andrew Gillen, Cheryl Carrico, David Knight, and Jake Grohs. Outcome expectations and environmental factors associated with engineering college-going: A case study. *Journal of precollege engineering education research*, 10(1), 2020.

[221] Holly M Matusovich, Ruth A Streveler, and Ronald L Miller. Why do students choose engineering? a qualitative, longitudinal investigation of students' motivational values. *Journal of Engineering Education*, 99(4):289–303, 2010. https://doi.org/10.1002/j.2168-9830.2010.tb01064.x.

[222] Margarita Mayo, Maria Kakarika, Juan Carlos Pastor, and Stéphane Brutus. Aligning or inflating your leadership self-image? a longitudinal study of responses to peer feedback in mba teams. *Academy of Management Learning & Education*, 11(4):631–652, 2012.

[223] Kristin McCue. Promotions and wage growth. *Journal of Labor Economics*, 14(2):175–209, 1996.

[224] Judith L Meece and M Gail Jones. Gender differences in motivation and strategy use in science: Are girls rote learners? *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 33(4):393–406, 1996.

[225] Meredith Meyer, Andrei Cimpian, and Sarah-Jane Leslie. Women are underrepresented in fields where success is believed to require brilliance. *Frontiers in Psychology*, 6:235, 2015.

[226] Katherine L Milkman, Modupe Akinola, and Dolly Chugh. What happens before? a field experiment exploring how pay and representation differentially shape bias on the pathway into organizations. *Journal of Applied Psychology*, 100(6):1678, 2015.

[227] David I Miller, Alice H Eagly, and Marcia C Linn. Women's representation in science predicts national gender-science stereotypes: Evidence from 66 nations. *Journal of Educational Psychology*, 107(3):631, 2015.

[228] Matthew K Miller, Julie P Martin, and Marisa K Orr. Toward determining changes in engineering-related social capital: resource composition as students make decisions about college. *Journal of Education and Training*, 1(2):72–91, 2014.

[229] Lawrence R Miller Jr. *Determining the factors that impact enrollment in cooperative education at the community college level*. PhD thesis, Walden University, 2018.

[230] Martin W Moakler Jr and Mikyong Minsun Kim. College major choice in stem: Revisiting confidence and demographic factors. *The Career Development Quarterly*, 62(2):128–142, 2014.

[231] Laurie A Morgan. Is engineering hostile to women? an analysis of data from the 1993 national survey of college graduates. *American Sociological Review*, 65(2):316–321, 2000.

[232] Corinne A Moss-Racusin, John F Dovidio, Victoria L Brescoll, Mark J Graham, and Jo Handelsman. Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, 109(41):16474–16479, 2012.

[233] Anna S Mueller, Tania M Jenkins, Melissa Osborne, Arjun Dayal, Daniel M O'Connor, and Vineet M Arora. Gender differences in attending physicians' feedback to residents: a qualitative analysis. *Journal of Graduate Medical Education*, 9(5):577–585, 2017.

[234] Sami Napari. Gender differences in early-career wage growth. *Labour Economics*, 16(2):140–148, 2009.

[235] Engineering Workforce Commission National Science Foundation. Engineering and technology enrollments, 2007.

[236] Urban Institute for Education Policy Center National Science Foundation. Summary report on the impact study of the national science foundation's program for women and girls, 2000.

[237] Jean Felix Ndashimye and Rajeev Darolia. Board 47: College engineering attainment among rural students. In *2019 ASEE Annual Conference & Exposition*, 2019.

[238] Julie A Nelson. Are women really more risk-averse than men? a re-analysis of the literature using expanded methods. *Journal of Economic Surveys*, 29(3):566–585, 2015.

[239] Virginie Nicaise, Julien E Bois, Stuart J Fairclough, Anthony J Amorose, and Geneviève Cogérino. Girls' and boys' perceptions of physical education teachers' feedback: Effects on performance and psychological responses. *Journal of sports sciences*, 25(8):915–926, 2007.

[240] Virginie Nicaise, Geneviève Cogérino, Julien Bois, and Anthony J Amorose. Students' perceptions of teacher feedback and physical competence in physical education classes: Gender effects. *Journal of teaching in Physical Education*, 25(1):36–57, 2006.

[241] OECD. Collaborative problem solving. (78), 2017. https://doi.org/https://doi.org/10.1787/cdae6d2e-en.

[242] Ontario Ministry of Education. Earning more credits through workplace experience. https://www.edu.gov.on.ca/morestudentsuccess/coop.html. Accessed: 30th June, 2017.

[243] National Academy of Engineering. Committee on Understanding the Engineering Education-Workforce Continuum. *Understanding the educational and career pathways of engineers*. National Academies Press, 2018.

[244] Canadian Council of Professional Engineers. Canadian engineers for tomorrow: Trends in engineering enrolment and degrees awarded 2001 to 2005. https://engineerscanada.ca/sites/all/themes/roots_book/lib/savetopdf.php?nid=5570, 2006. Accessed: 20th March, 2019.

[245] Matthew W Ohland, Marisa K Orr, Valerie Lundy-Wagner, Cindy P Veenstra, and Russell A Long. Chapter 8: Viewing access and persistence in engineering through a socioeconomic lens. *Engineering and social justice: In the university and beyond*, page 157, 2012.

[246] Craig A Olson and Brian E Becker. Sex discrimination in the promotion process. *ILR Review*, 36(4):624–641, 1983.

[247] Co operative Education and Canada Work-Integrated Learning. Wil definitions. https://cewilcanada.ca/CEWIL/CEWIL/About-Us/Work-Integrated-Learning.aspx, 2018. Accessed: 20th March, 2022.

[248] Marisa K Orr, Nichole M Ramirez, and Matthew W Ohland. Socioeconomic trends in engineering: Enrollment, persistence, and academic achievement. In *2011 ASEE Annual Conference & Exposition*, pages 22–1307, 2011.

[249] Marisa K Orr, Nichole M Ramirez, Matthew W Ohland, and Valerie Lundy-Wagner. Using high school and district economic variables to predict engineering persistence. In *2012 ASEE Annual Conference & Exposition*, pages 25–1427, 2012.

[250] Frank Pajares. Gender differences in mathematics self-efficacy beliefs. *Gender differences in mathematics: An integrative psychological approach*, pages 294–315, 2005.

[251] Grace Panther, Kacey Beddoes, and Cheryl Llewellyn. Salary negotiations and gender in engineering education. 2018.

[252] Gregory Park, David Lubinski, and Camilla P Benbow. Contrasting intellectual patterns predict creativity in the arts and sciences: Tracking intellectually precocious youth over 25 years. *Psychological Science*, 18(11):948–952, 2007.

[253] Mohammad S Parsa and Lukasz Golab. Social media mining to understand the impact of co-operative education on mental health. In *Proceedings of the International Conference on Educational Data Mining (EDM)*, pages 653–657. ERIC, 2020.

[254] Laura Parson. Are stem syllabi gendered? a feminist critical discourse analysis. *The Qualitative Report*, 21(1):102–116, 2016.

[255] Elizabeth A Paulin and Jennifer M Mellor. Gender, race, and promotions within a private-sector firm. *Industrial Relations: A Journal of Economy and Society*, 35(2):276–295, 1996.

[256] Alice L Pawley. Learning from small numbers: Studying ruling relations that gender and race the structure of us engineering education. *Journal of Engineering Education*, 108(1):13–31, 2019.

[257] Tuomas Pekkarinen and Juhana Vartiainen. Gender differences in promotion on a job ladder: Evidence from finnish metalworkers. *ILR Review*, 59(2):285–301, 2006.

[258] Michael R Pergamit and Jonathan R Veum. What is a promotion? *ILR Review*, 52(4):581–601, 1999.

[259] Andrea Perreault. Analysis of the distribution of gender in stem fields in canada. http://wiseatlantic.ca/wp-content/uploads/2018/03/WISEReport2017_final.pdf. Accessed: 20th March, 2019.

[260] Martin F Porter. Snowball: A language for stemming algorithms, 2001.

[261] Abigail Powell, Barbara Bagilhole, and Andrew Dainty. How women engineers do and undo gender: Consequences for gender equality. *Gender, work & organization*, 16(4):411–428, 2009.

[262] Stephen Pudney and Michael Shields. Gender, race, pay and promotion in the british nursing profession: estimation of a generalized ordered probit model. *Journal of Applied Econometrics*, 15(4):367–399, 2000.

[263] Joseph A Raelin, Margaret Bailey, Jerry Hamann, Leslie Pendleton, Jonathan Raelin, Rachelle Reisberg, and David Whitman. The effect of cooperative education on

change in self-efficacy among undergraduate students: Introducing work self-efficacy. *Journal of Cooperative Education and Internships*, 45(2):17–35, 2011.

[264] Edwin Ralph, Keith Walker, and Randy Wimmer. Practicum-education experiences: Post-interns' views. *The International journal of engineering education*, 25(1):122–130, 2009.

[265] Nichole Ramirez, Sally Smith, Colin Smith, Tessa Berg, Beata Strubel, Matthew Ohland, and Joyce Main. From interest to decision: A comparative exploration of student attitudes and pathways to co-op programs in the united states and the united kingdom. *International Journal of Engineering Education*, 32(5):1867–1878, 2016.

[266] Michael Ransom and Ronald L Oaxaca. Intrafirm mobility and sex differences in pay. *ILR Review*, 58(2):219–237, 2005.

[267] Erin D Reilly, Kadie R Rackley, and Germine H Awad. Perceptions of male and female stem aptitude: The moderating effect of benevolent and hostile sexism. *Journal of Career Development*, 44(2):159–173, 2017.

[268] Gusta Renes and Geert Ridder. Are women overqualified. *Labour Economics*, 2(1):3–18, 1995. https://doi.org/10.1016/0927-5371(95)80003-G.

[269] Ernesto Reuben, Paola Sapienza, and Luigi Zingales. How stereotypes impair women's careers in science. *Proceedings of the National Academy of Sciences*, page 201314788, 2014.

[270] Catherine Riegle-Crumb, Chelsea Moore, and Aida Ramos-Wada. Who wants to have a career in science or math? exploring adolescents' future aspirations by gender and race/ethnicity. *Science Education*, 95(3):458–476, 2011. https://doi.org/10.1002/sce.20431.

[271] Joseph Paul Robinson and Sarah Theule Lubienski. The development of gender achievement gaps in mathematics and reading during elementary and middle school: Examining direct cognitive assessments and teacher ratings. *American Educational Research Journal*, 48(2):268–302, 2011.

[272] Lisa Rosenthal, Bonita London, Sheri Robin Levy, and Marci Lobel. The roles of perceived identity compatibility and social support for women in a single-sex stem program at a co-educational university. *Sex Roles*, 65(9-10):725–736, 2011.

[273] Sue V Rosser. Attracting and retaining women in science and engineering. *Academe*, 89(4):24, 2003.

[274] Patricia M Rowe. Entry differences between students in cooperative and regular programs. *Journal of cooperative education*, 26(1):16–25, 1989.

[275] Sara L Rynes, Robert D Bretz Jr, and Barry Gerhart. The importance of recruitment in job choice: A different way of looking. *Personnel psychology*, 44(3):487–521, 1991.

[276] Mirka Saarela and Tommi Karkkainen. Discovering gender-specific knowledge from finnish basic education using PISA scale indices. In *Proc. of the Int. Conf. on Educational Data Mining (EDM)*, pages 60–67, 2014.

[277] Myra Sadker and David Sadker. *Failing at fairness: How America's schools cheat girls*. Simon and Schuster, 2010.

[278] Philip M Sadler, Gerhard Sonnert, Zahra Hazari, and Robert Tai. Stability and volatility of stem career interest in high school: A gender study. *Science education*, 96(3):411–427, 2012.

[279] Alpaslan Sahin, Adem Ekmekci, and Hersh C Waxman. The relationships among high school stem learning experiences, expectations, and mathematics and science efficacy and the likelihood of majoring in stem in college. *International Journal of Science Education*, 39(11):1549–1572, 2017.

[280] Anya Samek. Gender differences in job entry decisions: A university-wide field experiment. *Management Science*, 65(7):3272–3281, 2019. https://doi.org/10.1287/mnsc.2018.3107.

[281] Cate C Samuelson and Elizabeth Litzler. Community cultural wealth: An assets-based approach to persistence of engineering students of color. *Journal of Engineering Education*, 105(1):93–117, 2016. https://doi.org/10.1002/jee.20110.

[282] Karl W Sandberg. An exploratory study of women in micro enterprises: gender-related differences. *Journal of small business and enterprise development*, 10(4):408–417, 2003.

[283] Linda J Sax. Mathematical self-concept: How college reinforces the gender gap. *Research in Higher Education*, 35(2):141–166, 1994.

[284] Linda J Sax. Predicting gender and major-field differences in mathematical self-concept during college. *Journal of Women and Minorities in Science and Engineering*, 1(4), 1994.

[285] Carroll Seron, Susan S Silbey, Erin Cech, and Brian Rubineau. Persistence is cultural: Professional socialization and the reproduction of sex segregation. *Work and Occupations*, 43(2):178–214, 2016.

[286] Lily Shashaani. Gender differences in computer attitudes and use among college students. *Journal of Educational Computing Research*, 16(1):37–51, 1997.

[287] Jiyun Elizabeth L Shin, Sheri R Levy, and Bonita London. Effects of role model exposure on stem and non-stem student engagement. *Journal of Applied Social Psychology*, 46(7):410–427, 2016. https://doi.org/10.1111/jasp.12371.

[288] Rachel Silverman. Gender bias at work turns up in feedback, Sep 30 2015.

[289] Susan J Skledar, Barbara Martinelli, Kelley Wasicek, Scott Mark, and Robert J Weber. Training and recruiting future pharmacists through a hospital-based student internship program. *American Journal of Health-System Pharmacy*, 66(17):1560–1564, 2009.

[290] Adrienne Y. Smith. *They chose to major in engineering: A study of why women enter and persist in undergraduate engineering programs*. PhD thesis, 2012.

[291] David G Smith, Judith E Rosenstein, Margaret C Nikolov, and Darby A Chaney. The power of language: Gender, status, and agency in performance evaluations. *Sex Roles*, 80, 2019.

[292] Emma Smith and Stephen Gorard. Is there a shortage of scientists? a re-analysis of supply for the uk. *British Journal of Educational Studies*, 59(2):159–177, 2011. https://doi.org/10.1080/00071005.2011.578567.

[293] Kathleen N Smith and Joy Gaston Gayles. Girl power: Gendered academic and workplace experiences of college women in engineering. *Social Sciences*, 7(1):11, 2018.

[294] Kieran Snyder. The abrasiveness trap: High-achieving men and women are described differently in reviews. *Fortune Magazine*, 26:08–14, 2014.

[295] Matthew Soldner, Heather Rowan-Kenyon, Karen Kurotsuchi Inkelas, Jason Garvey, and Claire Robbins. Supporting students' intentions to persist in stem disciplines: The role of living-learning programs among other social-cognitive factors. *The Journal of Higher Education*, 83(3):311–336, 2012.

[296] Wei Song and Soon Cheol Park. A novel document clustering model based on latent semantic analysis. In *Third International Conference on Semantics, Knowledge and Grid (SKG 2007)*, pages 539–542. IEEE, 2007.

[297] Gerhard Sonnert, Mary Frank Fox, and Kristen Adkins. Undergraduate women in science and engineering: Effects of faculty, fields, and institutions over time. *Social Science Quarterly*, 88(5):1333–1356, 2007.

[298] Shaymaa E Sorour, Tsunenori Mine, Kazumasa Goda, and Sachio Hirokawa. Efficiency of lsa and k-means in predicting students' academic performance based on their comments data. In *CSEDU (1)*, pages 63–74, 2014.

[299] Elizabeth S Spelke. Sex differences in intrinsic aptitude for mathematics and science?: a critical review. *American Psychologist*, 60(9):950, 2005.

[300] Seymour Spilerman and Trond Petersen. Organizational structure, determinants of promotion, and gender differences in attainment. *Social Science Research*, 28(2):203–227, 1999.

[301] Jayne E Stake and Shannon D Nickens. Adolescent girls' and boys' science peer relationships and perceptions of the possible self as scientist. *Sex Roles*, 52(1-2):1–11, 2005.

[302] Jennifer Steele, Jacquelyn B James, and Rosalind Chait Barnett. Learning in a man's world: Examining the perceptions of undergraduate women in male-dominated academic areas. *Psychology of women quarterly*, 26(1):46–50, 2002.

[303] Rhea E Steinpreis, Katie A Anders, and Dawn Ritzke. The impact of gender on the review of the curricula vitae of job applicants and tenure candidates: A national empirical study. *Sex roles*, 41(7-8):509–528, 1999.

[304] Jacqueline Stevenson and Sue Clegg. Who cares? gender dynamics in the valuing of extra-curricular activities in higher education. *Gender and education*, 24(1):41–55, 2012.

[305] Jane G Stout, Nilanjana Dasgupta, Matthew Hunsinger, and Melissa A McManus. Steming the tide: using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (stem). *Journal of personality and social psychology*, 100(2):255, 2011.

[306] Beata Strubel, Joyce Main, Nichole Ramirez, Jake Davis, and Matthew Ohland. Modeling student perceived costs and benefits to cooperative education programs (co-ops) and pathways to participation. In *2015 IEEE Frontiers in Education Conference (FIE)*, pages 1–6. IEEE, 2015.

[307] Alexander Styhre, Maria Backman, and Sofia Börjesson. The gendered machine: concept car development at volvo car corporation. *Gender, Work & Organization*, 12(6):551–571, 2005.

[308] Rong Su, James Rounds, and Patrick Ian Armstrong. Men and things, women and people: a meta-analysis of sex differences in interests. *Psychological bulletin*, 135(6):859, 2009.

[309] Amanda Sullivan and Marina Umaschi Bers. Girls, boys, and bots: Gender differences in young children's performance on robotics and programming tasks. *Journal of Information Technology Education: Innovations in Practice*, 15:145–165, 2016.

[310] Radhika Suresh. The relationship between barrier courses and persistence in engineering. *Journal of College Student Retention: Research, Theory & Practice*, 8(2):215–239, 2006.

[311] IPEDS Completion Survey. American physical society. https://www.aps.org/programs/education/statistics/womenmajors.cfm. Accessed: 20th August, 2018.

[312] Robert H Tai, Christine Qi Liu, Adam V Maltese, and Xitao Fan. Planning early for careers in science. *Science*, 312(5777):1143–1144, 2006.

[313] Li Tan, Joyce B Main, and Rajeev Darolia. Using random forest analysis to identify student demographic and high school-level factors that predict college engineering major choice. *Journal of Engineering Education*, 110(3):572–593, 2021.

[314] Una Tellhed, Martin Bäckström, and Fredrik Björklund. Will i fit in and do well? the importance of social belongingness and self-efficacy for explaining gender differences in interest in stem and heed majors. *Sex roles*, 77(1):86–96, 2017. https://doi.org/10.1007/s11199-016-0694-y.

[315] Harriet R Tenenbaum. 'you'd be good at that': Gender patterns in parent-child talk about courses. *Social Development*, 18(2):447–463, 2009.

[316] Josh Terrell, Andrew Kofink, Justin Middleton, Clarissa Rainear, Emerson Murphy-Hill, Chris Parnin, and Jon Stallings. Gender differences and bias in open source: Pull request acceptance of women versus men. *PeerJ Computer Science*, 3:e111, 2017.

[317] Divy Thakkar, Nithya Sambasivan, Purva Kulkarni, Pratap Kalenahalli Sudarshan, and Kentaro Toyama. The unexpected entry and exodus of women in computing and hci in india. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 352. ACM, 2018.

[318] Glenn R Thiel and Nell T Hartley. Cooperative education: A natural synergy between business and academia. *SAM Advanced Management Journal*, 62(3):19–25, 1997.

[319] Heather Thiry. Why do students leave stem? *Bulletin of the American Physical Society*, 63, 2018.

[320] Joachim Tiedemann. Gender-related beliefs of teachers in elementary school mathematics. *Educational studies in Mathematics*, 41(2):191–207, 2000.

[321] Joachim Tiedemann. Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *Journal of Educational psychology*, 92(1):144, 2000.

[322] Andrew Toulis and Lukasz Golab. Graph mining to characterize competition for employment. In *Proceedings of the 2nd International Workshop on Network Data Analytics*, pages 1–7, 2017.

[323] Lisa Tsui. The effects of gender, education, and personal skills self-confidence on income in business management. *Sex Roles*, 38(5):363–373, 1998. https://doi.org/10.1023/a:1018701704103.

[324] National Center for Education Statistics U.S. Department of Education. 2003–04 beginning postsecondary students longitudinal study, second follow-up. http://nces.ed.gov/datalab/index.aspx, 2005. Accessed: 20th March, 2019.

[325] National Center for Education Statistics U.S. Department of Education. Higher education: gaps in access and persistence study. http://nces.ed.gov/pubs2012/2012046/index.asp, 2012. Accessed: 20th March, 2019.

[326] National Center for Education Statistics (NCES) U.S. Department of Education. Digest of education statistics. https://nces.ed.gov/programs/digest/2014menu_tables.asp, 2014.

[327] James R Valadez. Applying to college: Race, class, and gender differences. *Professional School Counseling*, 1(5):14–20, 1998.

[328] Virginia Valian. *Women at the Top in Science–And Elsewhere.* American Psychological Association, 2007.

[329] Jeffrey M Valla and Stephen J Ceci. Breadth-based models of women's underrepresentation in stem fields: An integrative commentary on schmidt (2011) and nye et al. (2012). *Perspectives on Psychological Science*, 9(2):219–224, 2014.

[330] Johanna Vennix, Perry den Brok, and Ruurd Taconis. Do outreach activities in secondary stem education motivate students and improve their attitudes towards stem? *International Journal of Science Education*, 40(11):1263–1283, 2018. https://doi.org/10.1080/09500693.2018.1473659.

[331] Beate Volker and Wouter Steenbeek. No evidence that gender contributes to personal research funding success in the netherlands: A reaction to van der lee and ellemers. *Proceedings of the National Academy of Sciences*, 112(51):E7036–E7037, 2015.

[332] Daniel Voyer and Susan D Voyer. Gender differences in scholastic achievement: A meta-analysis. *Psychological bulletin*, 140(4):1174, 2014.

[333] Jonathan Wai, David Lubinski, Camilla P Benbow, and James H Steiger. Accomplishment in science, technology, engineering, and mathematics (stem) and its relation to stem educational dose: A 25-year longitudinal study. *Journal of Educational Psychology*, 102(4):860, 2010.

[334] Jonathan Wai, Martha Putallaz, and Matthew C Makel. Studying intellectual outliers: Are there sex differences, and are the smart getting smarter? *Current Directions in Psychological Science*, 21(6):382–390, 2012.

[335] Gregory M Walton, Christine Logel, Jennifer M Peach, Steven J Spencer, and Mark P Zanna. Two brief interventions to mitigate a "chilly climate" transform women's experience, relationships, and achievement in engineering. *Journal of Educational Psychology*, 107(2):468, 2015.

[336] Ming-Te Wang. Educational and career interests in math: A longitudinal examination of the links between classroom environment, motivational beliefs, and interests. *Developmental psychology*, 48(6):1643, 2012. https://doi.org/10.1037/a0027247.

[337] Ming-Te Wang and Jessica Degol. Motivational pathways to stem career choices: Using expectancy–value perspective to understand individual and gender differences in stem fields. *Developmental Review*, 33(4):304–340, 2013. https://doi.org/10.1016/j.dr.2013.08.001.

[338] Ming-Te Wang, Jessica Degol, and Feifei Ye. Math achievement is important, but task values are critical, too: examining the intellectual and motivational factors leading to gender disparities in stem careers. *Frontiers in psychology*, 6:36, 2015.

[339] Ming-Te Wang and Jessica L. Degol. Gender gap in science, technology, engineering, and mathematics (stem): Current knowledge, implications for practice, policy, and future directions. *Educational Psychology Review*, 29(1):119–140, Mar 2017. https://doi.org/10.1007/s10648-015-9355-x.

[340] Ming-Te Wang, Jacquelynne S Eccles, and Sarah Kenny. Not lack of ability but more choice: Individual and gender differences in choice of careers in science, technology, engineering, and mathematics. *Psychological science*, 24(5):770–775, 2013.

[341] Ming-Te Wang, Feifei Ye, and Jessica Lauren Degol. Who chooses stem careers? using a relative cognitive strength and interest model to predict careers in science, technology, engineering, and mathematics. *Journal of youth and adolescence*, 46(8):1805–1820, 2017.

[342] Xueli Wang. Why students choose stem majors: Motivation, high school learning, and postsecondary context of support. *American Educational Research Journal*, 50(5):1081–1121, 2013.

[343] Jim Watters and Carmel Diezmann. Community partnerships for fostering student interest and engagement in stem. *Journal of STEM education*, 14(2):47–55, 2013.

[344] Erica A Wehrwein, Heidi L Lujan, and Stephen E DiCarlo. Gender differences in learning style preferences among undergraduate physiology students. *Advances in physiology education*, 2007. https://doi.org/10.1152/advan.00060.2006.

[345] Jeanna R Wieselmann, Emily A Dare, Elizabeth A Ring-Whalen, and Gillian H Roehrig. "i just do what the boys tell me": Exploring small group student interactions

in an integrated stem unit. *Journal of Research in Science Teaching*, 57(1):112–144, 2020. https://doi.org/10.1002/tea.21587.

[346] Joan Williams, Katherine W Phillips, and Erika V Hall. *Double jeopardy?: Gender bias against women of color in science*. Hastings College of the Law, Center for WorkLife Law, 2014.

[347] Melissa J Williams, Levy Paluck Elizabeth, and Julie Spencer-Rodgers. The masculinity of money: Automatic stereotypes predict gender differences in estimated salaries. *Psychology of Women Quarterly*, 34(1):7–20, 2010.

[348] Wendy M Williams and Stephen J Ceci. When scientists choose motherhood: a single factor goes a long way in explaining the dearth of women in math-intensive fields. how can we address it? *American scientist*, 100(2):138, 2012.

[349] Wendy M Williams and Stephen J Ceci. National hiring experiments reveal 2: 1 faculty preference for women on stem tenure track. *Proceedings of the National Academy of Sciences*, page 201418878, 2015.

[350] Rudolf Winter-Ebmer and Josef Zweimüller. Unequal assignment and unequal promotion in job ladders. *Journal of Labor Economics*, 15(1, Part 1):43–71, 1997.

[351] Shelly L Wismath and Maggie Zhong. Gender differences in university students'perceptions of and confidence in problem-solving abilities. *Journal of Women and Minorities in Science and Engineering*, 20(1), 2014.

[352] Rosalie Wyonch. Work-ready graduates: The role of co-op programs in labour market success. *CD Howe Institute Commentary*, 562, 2020.

[353] Vanessa L Wyss, Diane Heulskamp, and Cathy J Siebert. Increasing middle school student interest in stem careers with videos of scientists. *International Journal of Environmental and Science Education*, 7(4):501–522, 2012.

[354] Yonghong Jade Xu. Attrition of women in stem: Examining job/major congruence in the career choices of college graduates. *Journal of Career Development*, 44(1):3–19, 2017.

[355] Yi Xue and Richard C Larson. Stem crisis or stem surplus? yes and yes. *Monthly labor review*, 2015, 2015. https://doi.org/10.21916/mlr.2015.14.

Appendix

Appendix A

Table A1: Largest clusters of MECH job placements

Label	Words in cluster centroid	%	%ΔF
Project Management	projectmanag, mechan, arrang, written, verbal, timemanag, cad, assist, report, prepar	28%	7%***
Mechanical Design	solidwork, mechan, cad, manufactur, draw, assembl, machin, fabric, equip, prototyp	18%	-2%
Hardware Design	prototyp, hardwar, embed, electr, sensor, electron, c, pcb, circuit, devic	17%	0%
Web Development	web, javascript, framework, platform, mobil, languag, java, creat, passion, css	14%	2%
Manufacturing	manufactur, automot, mechan, equip, plant, supplier, processimprov, assembl, msoffic, improv	12%	-4%

Table A2: Largest clusters of Electrical job placements

Label	Words in cluster centroid	%	%ΔF
Application Development	passion, innov, app, idea, mobil, android, collabor, market, web, world	19%	-4%
Power Systems	electr, power, draw, autocad, safeti, shopdraw, prepar, equip, assist, load	17%	-1%
Quality Assurance	script, qa, autom, perl, execut, methodolog, java, qualiti, written, test	16%	0%
System Development	sql, net, c, web, databas, server, jqueri, framework, ui, mvc	12%	1%
Embedded Systems	circuit, pcb, schemat, electr, electron, hardwar, solder, analog, embed, sensor	9%	-1%
Hardware Design	fpga, analog, signal, digit, verif, pcb, vhdl, circuit, hardwar, schemat	8%	1%
Support Analyst	hardwar, xp, desktop, troubleshoot, remot, network, directori, instal, upgrad, customerservic	5%	5%
Integrated Circuits	matlab, camera, imag, sensor, electron, test, design, ccd, surveil, evalu	5%	2%
Control Systems	plc, electr, control, hmi, panel, equip, wire, robot, assembl, autocad	4%	3%
Transmission Systems	transmitt, sustain, ieso, climat, cleaner, foster, distributor, msaccess, adequaci, grid	3%	10%

Table A3: Largest clusters of Industrial job placements

Label	Words in cluster centroid	%	%ΔF
UI/UX	ux, user, passion, featur, app, mobil, creativ, css, impact, creat	24%	1%
Quality Assurance	analyst, script, qa, solut, strategi, agil, defect, enterpris, sql, concept	16%	-5%
Project Management	msoffic, powerpoint, analyst, msexcel, analyt, report, interperson, account, projectmanag, adhoc	15%	2%
Process Improvement	manufactur, mechan, caus, msoffic, root, chang, initi, processimprov, assist, equip	15%	7%
Software Development	css, javascript, html, jqueri, web, mysql, framework, api, databas, git	11%	-12%*
Supply Chain Analyst	incom, risk, decis, suppli, domest, oper, equiti, strategi, commit, benefit	8%	-1%
Financial Analyst	subsidiari, financi, wealth, invest, sharehold, analyst, casualti, insur, reinsur, agent	7%	5%

Table A4: Largest clusters of Chemical job placements

Label	Words in cluster centroid	%	%ΔF
Project Management	standard, account, msoffic, activ, support, schedul, cost, present, profici, packag	26%	7%
Manufacturing	plant, manufactur, equip, flow, chemic, processimprov, improv, elimin, instal, market	18%	-2%
Lab Assistant	advertis, arrang, chemic, lab, email, equip, sampl, materi, prepar, conduct	15%	9%
Energy	chemic, gas, mass, heat, pilot, plant, water, lab, energi, chemistri	15%	-16%**
Biotechnology	lab, assay, sampl, materi, microbiolog, chemic, french, formul, manufactur, biochem	10%	-15%*
Nanotechnology	chemic, synthesi, nanomateri, experiment, literatur, lab, research, polym, ftir, chemistri	7%	4%

Table A5: Largest clusters of Civil job placements

Label	Words in cluster centroid	%	%ΔF
Project Management	civil, construct, consult, autocad, structur, projectmanag, concret, contract, site, client	29%	7%
Capital Projects	civil, contract, construct, draw, tender, quantiti, survey, site, prepar, inspect	17%	-5%
Structural Design	advertis, arrang, draw, autocad, civil, construct, structur, site, bridg, draft	11%	-4%
Traffic Planning	traffic, road, synchro, licenc, sidewalk, citi, driver, municip, construct, civil	11%	-1%
Municipal Infrastructure	storm, sanitari, sewer, municip, water, watermain, civil, consult, construct, citi	7%	4%
Geotechnician	geotechn, soil, concret, asphalt, construct, inspect, survey, compact, driver, mine	7%	-19%*
Estimator	unsurpass, wait, belief, perspect, women, factor, spreadsheet, workforc, builder, aspir	5%	-17%
Rail Inspector	rail, inclement, ecolog, viabil, neighbourhood, fieldwork, pertin, conting, footprint, rehabilit	5%	8%
Diagnostics	forens, etab, masonri, draft, wood, reclad, aid, sap2000, structur, facad	4%	4%

Table A6: Largest clusters of Nanotechnology job placements

Label	Words in cluster centroid	%	%ΔF
Software Development	collabor, problemsolv, databas, function, busi, sql, data, writtencomm, assess, tool	27%	-3%
Lab Assistant	advertis, arrang, lab, email, research, cellulos, scientif, chemic, data, write	15%	8%
Manufacturing	manufactur, safeti, materi, suppli, chemic, equip, prepar, plant, produc, sampl	14%	-2%
Optics	optic, experiment, sequenc, randd, matlab, remind, lab, advisor, discuss, scientist	14%	9%
Electronics	electron, matlab, microfluid, chemistri, lab, research, fabric, circuit, devic, microscopi	13%	-7%
Materials	synthesi, lab, experiment, chemic, polym, synthes, nanoparticl, research, chromatographi, arrang	11%	4%

Table A7: Largest clusters of Environment job placements

Label	Words in cluster centroid	%	%ΔF
Research Assistant	environment, research, present, assess, conduct, studi, water, assist, lab, prepar	35%	11%*
Civil Design	survey, civil, construct, draw, municip, prepar, autocad, contract, road, inspect	15%	-15%*
Water Infrastructure	water, sewer, sanitari, civil, hydrolog, storm, municip, hydraul, consult, environment	11%	7%
Survey and Design Jobs (arranged by student)	advertis, arrang, survey, autocad, environment, draw, famili, instrument, calcul, friend	9%	1%
Hydrogeology	groundwat, unrestrict, licenc, sampl, water, soil, driver, hydrogeolog, vehicl, environment	9%	-2%
Geologist	geolog, geotechn, mine, soil, slope, environment, stabil, consult, investig, tail	9%	-5%
Geotechnician	geotechn, soil, asphalt, compact, concret, hydrogeolog, drill, sampl, terraprob, technician	7%	-30%**

Table A8: Largest clusters of Biomedical job placements

Label	Words in cluster centroid	%	%ΔF
Systems Analyst	iso, medic, patient, compliant, instruct, standard, regul, document, iec, chang	29%	-2%
Clinical Analyst	priorit, cure, clinic, societi, written, propos, accompani, assess, communitech, research	25%	15%
Research Assistant	biomed, medic, advertis, arrang, lab, biolog, matlab, research, literatur, scientist	23%	-4%
System Development	c++, script, content, media, c, platform, geek, wordpress, desktop, summari	17%	8%