

Reddit Mining to Understand Women’s Issues in STEM

Allison Jacobs
University of Waterloo
Waterloo, Ontario, Canada
adjacobs@uwaterloo.ca

Shivangi Chopra
University of Waterloo
Waterloo, Ontario, Canada
s9chopra@uwaterloo.ca

Lukasz Golab
University of Waterloo
Waterloo, Ontario, Canada
lgolab@uwaterloo.ca

ABSTRACT

We analyze social media discussions to understand the issues experienced by women in Science, Technology, Mathematics and Engineering (STEM), both in the classroom and at work. To do so, we identify two relevant discussion forums on the Reddit social media platform: r/womenEngineers and r/xxSTEM, and we perform topic modelling to discover frequently discussed topics. Our analysis leads to three main findings. First, women discuss harassment, inequality, and lack of representation. These issues were found to persist in both academic and professional environments. Second, women express a lack of confidence and feelings of being an imposter as well as having to change their appearance and behaviour to fit in with their coworkers. These issues again did not disappear after academic life but tended to persist into professional careers. Finally, some women were faced with a lack of parental leave policies at their organizations.

Keywords: Women in Engineering, STEM, Social Media Mining, Text Mining, Gender Imbalance

1 INTRODUCTION

In recent years, marketing various Science, Technology, Engineering and Mathematics (STEM) programs to young women has become important to post-secondary institutions in an attempt to even the gender gap [2]. As a result, the percentage of women enrolled and graduating from engineering programs has seen improvement over the past decade [23]. That said, the proportion of female graduates earning bachelor’s degrees in fields such as computer science and engineering is still low (around 20%) [23, 51]. Although the attraction to engineering and the success rates from an academic standpoint are increasing, there is a large percentage of women who do not pursue a career in engineering or leave the workforce [1]. For example, a longitudinal study with 87 men and 34 women chemical engineering undergraduates found that women scored better than men on pre-college admission tests and were highly motivated at the start of their program, but, by the end of their sophomore year, twice the percentage of women, in comparison to men, dropped out [21]. Studies also suggest that many female engineering graduates switch out of the field right after graduation or sometime after joining the workforce [27]. Overall, women make up less than 14 percent of the professional engineering workforce in Canada, a number that has remained constant for some time [8].

Prior work has studied gender differences in STEM academics and careers from various perspectives, including differences in biology, aptitude, attitude, motivation, interests, performance, choice, socio-cultural factors, and opportunity [55]. With a focus on women’s barriers to entry and attrition [18, 49], studies have

listed issues such as dissatisfaction over pay and promotion opportunities [27], self-doubt [17, 18, 48], lifestyle values of raising a family and altruism [52, 55], under-representation [10], and sexism [9, 18]. However, most of this research is based on samples collected during interviews and is mostly qualitative.

In the context of data-intensive work on gender issues, there is work that analyzed gender differences in undergraduate Engineering applicants [11] and work-integrated learning programs [12, 13], and social media mining studies of the #MeToo [30, 37, 43] and the #ILookLikeAnEngineer [28] movements. With social media providing an anonymous environment to discuss sensitive topics [54], it has been used to analyze issues such as suicide [26], domestic violence [46, 47], mental health [15, 31, 41], bullying [58], and racial inequality [16]. However, we are not aware of any social media mining studies that analyze women’s issues in STEM.

To fill this gap, we contribute to the research on gender issues in STEM by analyzing the Reddit social media platform. We identified two Reddit discussion forums – r/womenEngineers and r/xxSTEM – designed for women to share their experiences in STEM academics and workforce, and to offer advice to others considering or newly entering the STEM workforce. Specifically, the goal of this analysis is to identify topics that are frequently discussed by women in STEM. To do so, we perform *topic modelling* of the content of the above Reddit forums between December 2015 and February 2019, using Non-negative Matrix Factorization (NMF) [40].

Our analysis leads to three main findings. First, women discuss harassment, inequality, and lack of representation. These issues were found to persist in both academic and professional environments. Second, women express a lack of confidence and feelings of being an imposter as well as having to change their appearance and behaviour to fit in with their coworkers. These issues also did not disappear after academic life but persisted into professional careers. Finally, some women were faced with a lack of parental leave policies at their organizations. Our social media analysis confirms the findings obtained by qualitative prior work and provides more specific details into the barriers faced by women in STEM.

The remainder of this paper is organized as follows. Section 2 explains our dataset and method; Section 3 presents the results; Section 4 discusses our findings and compares them to prior work; and Section 5 concludes the paper with directions for future work.

2 DATA AND METHOD

2.1 Data

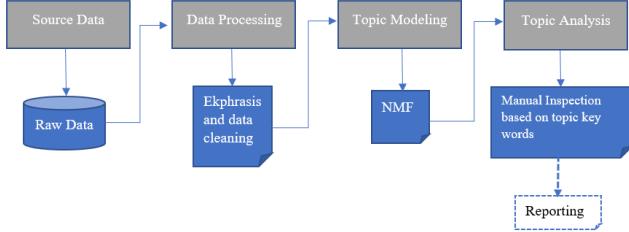
Reddit is a social media platform divided into discussion forums referred to as *subreddits*. A subreddit contains *posts* that initiate a discussion, and *comments* on the posts. This analysis is based on Reddit data from December 2015 to February 2019, downloaded from a publicly available database stored on Google Big Query¹.

© 2020 Copyright for this paper by its author(s). Published in the Workshop Proceedings of the EDBT/ICDT 2020 Joint Conference (March 30-April 2, 2020, Copenhagen, Denmark) on CEUR-WS.org. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0)

¹<https://cloud.google.com/bigquery/>

Table 1: Subreddit Statistics

	r/womenEngineers	r/xxSTEM
Number of Posts	243	702
Number of Comments	1041	2305
Number of Unique Users	388	763
Number of One-Time Users	205	413

**Figure 1: Summary of Method**

There are over one million subreddits in total. To find those which are most closely related to our study of issues experienced by women in STEM, we first identified all subreddits containing at least one of the following terms in their title: “women”, “engineer”, “STEM”, or “feminism”. We selected these words to capture elements of STEM, specifically engineering, as well as feminism. We then read a sample of recent posts in each of these subreddits to determine the topics being discussed among posts as well as common themes in the discussions. We selected two subreddits for analysis — r/womenEngineers and r/xxSTEM — that included discussions of gender issues in STEM education and workplace. Examples of subreddits that we examined but discarded are r/Feminism (containing discussions of gender issues, but not related to STEM), r/AskEngineers (containing questions about engineering education and careers, but not from a gender perspective), and r/AskWomen (containing women’s perspective on various issues, but not related to STEM).

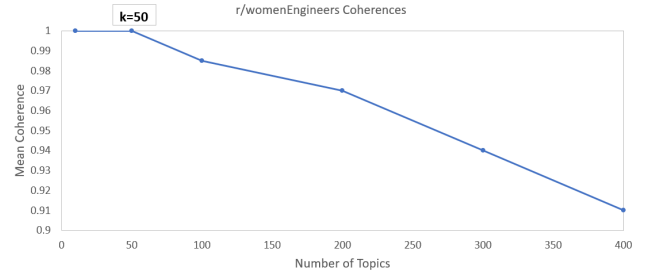
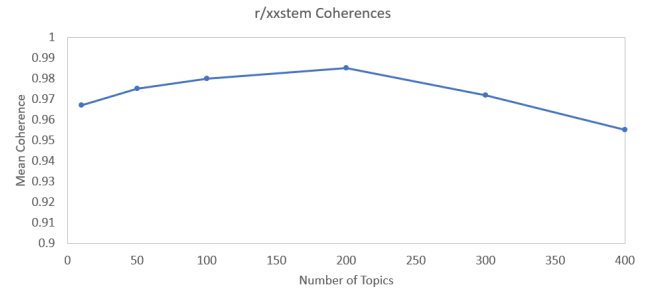
Table 1 contains some statistics for each of the chosen subreddits. The number of unique users is the number of users who created at least one post or comment on the forum, whereas one-time users are those who only created one post or comment. Over half the users on both forums are one-time users, a trend previously observed in social media by [15, 41]. Notably, the number of users who posted or commented at least once on both forums is only 31, suggesting that the content on r/womenEngineers and r/xxSTEM may be different, and thus motivating the need to analyze both of these forums.

We inspected a sample of posts and comments from the two selected subreddits. We observed that a post starts a conversation and the sharing of experiences, but the comments on each post provide more insight into experiences and issues faced. Additionally, there are more comments than posts in total. As a result, we selected only the comments for further analysis.

2.2 Method

Our method to analyze the comments made on r/womenEngineers and r/xxSTEM is summarized in Figure 1 and detailed below.

Data pre-processing: We discarded comments of the form “[REMOVED]” and “[DELETED]”, indicating that they were removed by Reddit moderators. For the remaining comments, we used Ekphrasis [3] to parse them, normalize the words, remove stopwords, and correct spelling errors. There were originally

**Figure 2: Coherence Plot for different numbers of topics; r/womenEngineers****Figure 3: Coherence Plot for different numbers of topics; r/xxSTEM**

1095 comments in r/womenEngineers and 2446 in r/xxSTEM. Discarding the deleted comments, we are left with 1041 and 2305, respectively, as reported in Table 1.

Topic Modeling: Next, we run the Non-negative Matrix Factorization (NMF) algorithm to segment the comments and identify underlying topics. NMF groups the comments into k clusters, with k being a user-specified parameter, and identifies a set of representative terms, which we refer to as *topic keywords*, for each cluster.

To decide on the number of clusters or topics, we experimented with $k \in \{10, 50, 100, 200, 300, 400\}$, and computed the coherence of each choice [38]. Topic coherence allows for analysis to determine the extent to which the top terms representing a topic are semantically related, relative to some “background corpus”. A higher coherence is desirable.

We plot the coherence as a function of the number of topics in r/womenEngineers and r/xxSTEM in Figure 2 and Figure 3, respectively. The y-axes of the figures start at 0.9, not 0, to emphasize the differences in coherence. For r/womenEngineers, $k = 50$ has the highest coherence. For r/xxSTEM, highest coherence occurred at 200 topics, but the coherence at 50 topics was nearly as high. Standard data science practice dictates an affinity for a smaller model, and to maintain consistency across subreddits, we chose $k = 50$ for both subreddits.

To verify the choice of k , we manually inspected the representative terms of each cluster for each tested value of k . For $k < 50$, NMF returned terms that suggested several topics per cluster, whereas larger values of k produced clusters of repeating topics.

Topic Analysis: In this step, we manually examined the keywords returned by NMF for each topic, and we analyzed the comments belonging to each topic. Finally, after inspecting the topic keywords and reading a sample of comments assigned

to each topic, we manually extracted five frequently discussed *themes* that will be discussed below.

3 RESULTS

3.1 Topic Analysis

Table 2 and Table 3 show the topic keywords produced by NMF for five largest topics for the two subreddits, as well as the number of comments categorized under each of these topics. For r/womenEngineers, the terms suggest frequent discussions of classroom experiences, job interview experiences, parental leave, and pay. For r/xxSTEM, the representative terms suggest discussions of gender bias and gender gap.

We identified the following five *themes* that frequently occurred in the discussions on both subreddits. Notably, these five themes were present both in discussions of academic and workplace issues.

- (1) Representation: discussions of the lack of representation of women.
- (2) Personal: an experience that was unique to the person; e.g., an encounter that was sexist or classified as harassment.
- (3) Confidence: imposter syndrome or other internal battles, i.e., feelings of inadequacy or self-doubt.
- (4) Interests: expressions of interests that have allowed women to continue or discontinue their engineering career.
- (5) Equality: comparisons to another gender or person.

3.2 Analysis of r/womenEngineers

We now discuss the above five themes in more detail, starting with r/womenEngineers. Table 4 summarizes the types of issues discussed under each theme in the context of academics and careers.

In terms of academics, we found Equality to be frequently discussed. This includes women comparing themselves to their male peers and discussing experiences in which they were treated differently. We give an example comment below.

Equality: *"I recently worked with a male project partner on an electronics project for grad school, he didn't know how to hook up the multimeter to measure current. Was that failure of knowledge because he was male? Nope, he just didn't have much hands-on experience with hardware. I think a lot of women growing up do really well in school (which is often theory heavy) but aren't necessarily offered the practical experience outside of school. Men are assumed to like tools and technology from a young age so it's more likely their family members give them gifts that align with those things and invite them to talk about and participate in engineering related projects. So, it can seem like women aren't as good but the actual reason isn't gender but societal influence."*

The Personal theme was also frequent. In these comments, women were recounting experiences of harassment from their school life.

Personal: *"Guy in charge of the school machine shop who sets all the equipment for his size at 6'3" where I am 5'3": why are you on your tip toes, you cut like a girl (how does one cut like a girl on a band saw?)"*

Next, the topic of Representation included comments discussing the lack of female classmates, faculty members, and role

models.

Representation: *"I went to school in a mostly male major, and I'm currently working in automotive manufacturing. It's great having male friends, but every now and then I just want to be able to sit down and talk about nail polish colors with other girls. It's good for your health, plus there's a lot of conversations that are girls-only. My company is really small, so there's not too many women I can hang out with after work. But it makes it so much more worth while when I get that little break and I hangout with other girls!"*

Confidence discussions frequented around the theme of imposter syndrome and self-doubt. Women discussed the guilt they felt for receiving scholarships and in general any form of academic success.

Confidence: *"That being said, I have also suffered from "Imposter Syndrome" a lot recently, so I completely understand where you are coming from. It's intimidating being in a field dominated by men where you basically stick out like a sore thumb. All you can do is your best and seek feedback/guidance whenever possible."*

Finally, Interests included discussions about the impact of engineering on the community as well as the reasons why the individual has chosen to pursue an engineering education.

Interests: *"I was pretty lucky. Several adults recognized my aptitude for engineering before I knew engineering was a thing. I was encouraged by my dad and several teachers to continue developing my skills in math and science, because I always excelled in those. Once I realized that I could use those skills (which I also genuinely enjoyed) to help contribute to new technology, I felt it was a perfect fit. I always had a passion for learning, so I wanted to be at the vanguard, learning all the time."*

At the career level, many comments referred to Personal experiences. This includes sharing experiences of sexism, harassment or other uncomfortable encounters. We show an example below.

Personal: *"I know it sucks to hear, but you need to grow tougher skin and stand your ground. You need to pull that guy aside and frankly tell him that you're not interested. If he continues to harass you, that is where you would consider going to the Title IX department. Continuous harassment being saying inappropriate things to you, gossiping about you behind your back, or sending you sexually explicit images/messages. I had to go there to file a claim against a GTA who blatantly sexually harassed me and they're nothing but welcoming."*

For the theme of Equality, women compared themselves to their male coworkers and commented on the wage gap.

Equality: *"I was in the same boat as you. I was paid way less than my male colleagues had more education - master's degree vs colleagues having bachelor's or some no degree. I was the only female engineer and other female engineers who started in my department were moved to non technical positions. (No idea if that is what they wanted just something I noticed) I was given the soft jobs, less technical and after over a year of asking for the more interesting technical jobs, I ended up applying to other jobs out of the company and left for job that paid me 20% more and for more*

Table 2: Five largest topics: r/womenEngineers

Topic keywords	Number of comments
Topic 47: go would make thing also say school ask differ experi lot find woman someth men class talk much could student well someone like sure back work male start think engin	218
Topic 0: like feel general seem person thing someth respons interview scienc right black field sometim part understand life crap yeah uncomfot awkward fashion may gender expect math unfair non shadow struggl	40
Topic 3: number year graduat school go age start semest senior program week percent kid leav hour month never back gone bachelor three worri junior money pay employe class next long	37
Topic 2: engin degre field mechan comput studi mani life mind electr industri scienc sub probabl area graduat hey better tech male see chang thread relat societi live question lot design role	31
Topic 33: wear shirt casual jean offic blous dress slack black pant heel cardigan button comfort nice shoe color flat hair occasion cloth busi manufactur boot usual feminin tend fashion coupl tech	29

Table 3: Five largest topics: r/xxSTEM

Topic keywords	Number of comments
Topic 46: think want need way ask also could see even someth talk differ take person someone experi find use question point start right lot guy give meet never may team happen	382
Topic 7: women men stem gender issu confer woman bias male field tech societi support group mani secur hire career us young think lot equal gap reason problem posit cultur treat event	83
Topic 9: week last start next project today back bridg pe hour plan review trip still weekend busi big finish coupl current already excit hous glad super new fire home friday laterg	79
Topic 42: program degre phd scienc year appli comput internship univers requir start master interest graduat cs math stem also lab student mani data undergradu languag research applic offer cours field colleg	76
Topic 23: school grad student high undergrad gpa appli depart physic stress math colleg univers year lot went back choic learn first late graduat still strong professor anxieti guy start top major	59

Table 4: Discussion themes in r/womenEngineers

	Representation	Personal	Confidence	Interests	Equality
Academics	lack of female classmates, comparing enrollment rates	harassment from peers and professors	imposter syndrome in program, guilt for scholarships, guilt for academic success, comparing skills to male peers	family implications, applications to industry	comparing to male peers, differences in promoting STEM programs
Career	lack of female coworkers, lack of mentors	harassment from coworkers and managers	imposter syndrome in position and worth in company	starting a family, work impact, changing perspectives and interests	comparing to male coworkers, wage gaps, unequal treatment

technical and interesting work. And I'm good at it :)"

Representation again discusses the lack of female coworkers or what it feels like to be the first female employee hired at a company. One comment example is below.

Representation: *"If you aren't able to find that within your own organization, look to similar ones. At my last job I was the only woman in our entire department, and the only women more senior than myself were in HR and Marketing (not that I was that senior, it's just that I worked at a brewery and there just aren't a lot of women). So I attended professional events like SWE and also industry specific events, and found women who were in those higher level positions but with different companies. I also met women just*

through the vendors and contractors I worked alongside. One of the biggest mentors I had was the COO of one of our vendor companies, and I met her because she happened to come on site one day to see how our installation and testing was going."

Confidence discussions again related to the imposter syndrome and understanding one's worth within the company.

Confidence: *"The larger point to understand here is how the system being tilted in your favor doesn't do you any favors. Now a feeling of imposter-syndrome goes from being an anxiety-induced-fabrication to justified and you can't really tell how much of it is justified vs. emotional until you're on the job and start accomplishing tasks. Because the tilt of the system has pulled you ahead*

of yourself you cannot perform. Because you cannot perform you cannot be trusted with critical tasks. Because you are not trusted with worthwhile tasks you do not improve."

In a career context, discussions of Interests included reasons why women decided on their careers, and reasons for leaving a job. Furthermore, commenters discussed the consequences associated with deciding to start a family and the potential glass ceiling issues associated with promotions and career growth.

Interests: *"The worst issue was at my last job when my boss and I sat down to discuss my "5-year plan" and mentioned how I may want to get switch jobs due to the heavy travel in a few years when I meet a man and want to settle down. I told him nope that I love my job and don't plan to leave because of the travel and possible children (which is light years away, if at all). He said well, we'll keep it in mind that you might want to slow down."*

3.3 Analysis of r/xxSTEM

This analysis was repeated in the same manner for the second subreddit, r/xxSTEM. Table 5 summarizes the academic and career issues discussed on this forum.

For Academic issues, many discussions were Personal. This included recounts of comparisons and assumptions made about a student based on their gender or recounts of harassment. We give an example comment below.

Personal: *"Of course you got a job at [company]. You're a girl." - classmate. "Girls get all the scholarships." - classmate "Being a woman must have helped you get the job." - my FEMALE co-op coordinator "I never hire women. All they do is get pregnant and go on maternity leave." - my sister's boss*

Discussions of Equality described how to be seen and not treated differently from male peers. An example of this is below.

Equality: *"In my high school, girls were actively discouraged/not allowed to take these classes. I wanted to take wood shop, but wasn't allowed. I ended up doing theatre and learned to build sets instead."*

One commenter discusses her early exposure to STEM and its impact.

Equality: *"I was lucky my parents did this while I was growing up. I helped my Dad with his construction projects, he helped my sibling and I with our homework (math & sci included), he took us to his work and other future engineering events, we had legos and K'nex, and his friends/coworkers were fairly diverse engineers. I was not aware that women were an oddity in STEM until late in highschool"*

Representation discussions followed similar themes present in r/womenEngineers.

Representation: *"Of the 104 people who are "members of the department" (PhD candidates, instructors, professors, research faculty, and post docs) only 17 are women. And 9 of those 17 are professors or instructors and thus have keys to private restrooms. It just pisses me off that I'm so non-existent to the department that I have to go down 4 floors for a bathroom when my office and lab are up here. I'm an undergrad at an R1 school with 40,000 people."*

Interests focused on the difficulties starting a family in academia.

Interests: *"The thing that really annoys me is that academia in general is terrible at taking care of parents, especially mothers. My graduate school didn't have a maternity leave policy in place until two years ago. They presented that policy at the school wellness meeting and it was SO embarrassing. But even that maternity leave policy is terrible. Only 4 weeks paid leave and that's only for the mother. Anyways this is a long way of saying I agree with you and I think the above stuff is a lot of the reason why women get lost in the academia pipeline. They are basically actively punished for wanting a family."*

Confidence was commonly challenged for female students, which has led them to feelings of self-doubt.

Confidence: *"How do you know you got accepted purely on merit and not because you are a woman filling their quota?" Said to me by a woman. I don't think she knows that 3/4 of my class and every class below me has been female. We're not filling a quota, apparently women have stronger applications on average."*

At the career level, Equality and Personal issues made up most of the comments. Below is an example of a comment discussing Equality.

Equality: *"Women are consistently held to an impossibly higher standard compared to men in life in general it seems. So, if your experience in life is that you have to be pretty damn near perfect, you're not going apply to something that you know you won't get if you don't meet all the qualifications for. Whilst your male colleagues, who society just automatically gives more respect to, can afford to apply to things without all the qualifications."*

An example of a positive comment on equality is shown below.

Equality: *"Yes it does! But not everywhere. The big tech companies are great and actively work to make women and other minority groups feel welcome. I had a similar experience to you in school, constantly feeling like an outsider. After a few years of work and picking the right companies, I'm happy to say I feel accepted as an engineer— not just a female engineer. Echoing a commenter above, don't forget to advocate for yourself. You are your own best cheerleader and won't be happy in any situation if you let people push you around."*

Another commenter further discusses comments made to them about wage gaps and other general comparisons to their male coworkers.

Equality: " 1. "Have fun making .75 for every dollar a man makes." 2. "You're only here to get your Mrs. degree." Which is a degree made up to imply that I'm only in college to find a husband to support me. I am currently on a web development team and my supervisor is female, but the rest of my team of 2 besides me are male. It is so important for more women to be in the engineering/computer science field. Best of luck, engineering is tough but it's worth it!"

Personal experiences including sexism were common as well as general gender bias.

Table 5: Discussion themes in r/xxSTEM

	Representation	Personal	Confidence	Interests	Equality
Academics	lack of female faculty, lack of female research peers, minority in classroom environment	harassment from peers and professors, gender bias and assumptions	imposter syndrome, not standing up for oneself	why STEM is chosen, future career goals, families within academia	comparing to male peers, STEM encouragement differences
Career	lack of female coworkers, lack of other minorities and mentors in the workplace, difficulty finding someone like themselves	harassment from coworkers, bias against gender, gender assumptions	imposter syndrome, talking and presenting with confidence to coworkers, how to be taken seriously	personal goals and interests, family implications	comparing to male coworkers, wage gaps, gender gaps, mistreatment

Personal: *"I never reported my harassers. Other women at the company reported their harassers and those women never worked there much longer. I wish I had been brave enough to do it but I saw the pattern of report, quit or be fired and decided I wasn't ready to leave yet. I was the only woman on the team for a long while. I left the job this past August and moved to a way less hostile environment."*

Although women know there are a lack of role models in their fields, there are those who are determined to see this change. One comment reads:

Representation: *"I came here to say this, a lot is based in gender expectations and a lack of inspirational role models, you rarely see women in these roles in the media and that needs to be challenged and improved. It's one of the reasons that I do outreach. It annoys me but I know that by being there I am showing that women can and do have jobs like this."*

A recurring theme in the discussion of Interests was again focused on maternity leave policies and how this impacts the choice of company in the future.

Interests: *"The culture will vary company to company. I would recommend reaching out to mothers/women you can identify on LinkedIn or through other networking and contact them directly to ask about this info. This is the kind of thing where you really do have to talk to employees outside the interview context to understand the company culture and how that influences use of maternity leave policies."*

Finally discussions of Confidence again revolve around self-doubt.

Confidence: *"We have a weekly status meeting as a team, and I always make sure to talk about 2-3 things I did. The problem is if someone else talks about something before I get a chance, there's no need to repeat what's already been said. I definitely have a confidence problem. I still feel really unsure of what I'm doing, which is why I usually seek confirmation from someone else. I think I'm really afraid to mess up or seem like I don't know what I'm doing."*

4 DISCUSSION AND RELATED WORK

In general, this analysis revealed negative results. One reason for the lack of positive discussions is that anonymous social media platforms often provide a way to vent, or discuss problems and invite guidance and support from those who have experienced similar situations.

Below, we reiterate our main findings, organized around our five themes, and compare them to those reported in prior work.

Confidence: Women shared their problems with confidence in their academics and careers or experienced some type of imposter syndrome and self-doubt. Past studies support this finding and report that women start college with a lower self-confidence [57], a lower self-concept about their performance in STEM courses [44, 45], and a lower self-efficacy [34, 39]. As students progressed through the curriculum, bad grades led to greater drops in self-esteem for women in engineering [14]. Additionally, Fouad et. al found that women who leave engineering have lower levels of self-efficacy than women who stay in the profession [36]. Other studies also identify self-doubt as an important reason for women's attrition from STEM [17, 18, 48].

Representation: We found that support and mentorship are critical to combating the confidence issues women may face. Women noted that mentorship was difficult to find in a male-dominated workforce. We found that women were frequently asking for resources in which they can find the support that is lacking in the workplace environment. Past studies support this finding and suggest that under-representation may discourage females to pursue or continue in STEM [10, 50]. Similarly, Cadaret et al. suggest that career interventions can assist young women to strengthen their expectations for success in engineering fields, thus expanding their social support and improving their identity as engineers through effective mentoring [33].

Further to this, we found that improper forms of education counselling may prevent women from pursuing a STEM related career. Studies show that a female student's choice not to pursue a career in STEM is not due to their lack of technical skills, but rather the perception of a lack of skills by parents, counsellors and students themselves [5, 53, 55]. Textbooks mentioning the work and photographs of male scientists propagate a similar gender bias [4], emphasizing the social and occupational segregation of gender [32].

Interests: In the discussions of interests, we note that women discussed the implications of starting a family while also trying to succeed in their career. Studies have noted similar circumstances

affecting women's careers, such as raising children, accommodating a spouse's career, or caring for elderly family members [7, 22, 52, 55]. Women also described cases of discrimination and poor infrastructure in their workplace as they may have been the only employee pregnant at a time or one of a limited number of female employees. One civil engineer recounts her experience by noting that she was the first female her employer has ever hired, and that the maternity policy was never needed before her. It was also noted that although Information Technology related industries are implementing maternity policies, industries such as construction are failing to change to meet the needs of their employees, specifically women. This underscores the need for clear parental leave information dissemination to all employees. Wang and Degol note similar issues in STEM workplaces which often lack the support required by women with young children and other care-taking responsibilities, which forces them to vacate STEM positions at greater rates than men [55].

Other posts noted that starting a family would result in missed opportunities and chances to be promoted. Their managers suggested they take jobs with less responsibility and no travel, so they can be ready for family life. This disparity was discussed by Hewlett [25]. After conducting a survey to explore the professional and private lives of highly educated, high-earning professionals, she found that 49 percent of those women were childless compared to only 19 percent of men. Furthermore, she discusses the wage gap and its connection to the penalties women tend to face when they interrupt their careers to have children. Recent studies have confirmed that an increasingly large part of the wage gap can now be explained by childbearing as this is something that interrupts a woman's career, but not a man's [6, 24].

Equality: Our results indicate that some women were treated unequally by their classmates, coworkers, faculty, and managers. Differences in terms of career advice, kind of work assigned, wage, and promotion, have been noted. Past studies found similar instances of inequality. Researchers found that male faculty were more likely to ignore emails from prospective female graduate students than males [35]. Analysis of 1,224 recommendation letters for graduate studies in geo-science revealed that female applicants were half as likely to receive excellent versus good letters compared to male applicants. Male and female recommenders were equally likely to display this bias [19]. In STEM workplaces, women's success in STEM was attributed to luck as opposed to skill [56]. Other studies conclude that men have a higher chance of being promoted and receive higher increases in wage throughout their careers [20, 29].

Personal: Our results indicate that women feel pressured to change themselves to fit in with their male colleagues and peers. This includes providing advice on clothing that would allow women to be taken seriously, and avoiding unwelcome harassment or interactions. Past studies support this finding [42, 56]. Some women have noted that they were forced to change themselves in some way to meet the culture in their workplace or in an academic setting. This includes avoiding reporting their harassers out of worry of further implications or consequences. From a young age, women have noted the differences in treatment and general gender disparities. For some, STEM careers were not promoted to them due to their gender. In other cases, those who decided to pursue their aspirations were anomalies who then encountered various issues in school and in the workplace. Past studies note similar instances of implicit and explicit sexism in both STEM academics and workplaces [18, 48].

5 CONCLUSIONS

In our analysis of the subreddits r/womenEngineers and r/xxSTEM, we have found that there is valuable information being discussed by women at various points in their STEM careers. Beginning at an academic level and proceeding into their professional career, women have noted different reasons for dissatisfaction and attrition. A trend that we were not expecting to see was the many discussions of beginning a family while also being a woman in engineering. Unequal treatment from male coworkers as well as employers not having a maternity or paternity policy in place have also placed a strain on women.

These findings motivate the need for policies to effectively support working mothers and allow them to succeed in their career without dealing with a glass ceiling mentality and sacrificing the want for a family to advance professionally. This issue combines both the lack of equality in the workforce and the lack of representation. Women are commonly faced with the decision between their career and a family and are therefore unable to move up the corporate ladder. They cannot situate themselves in a position where they are able to become mentors and role models for the upcoming generation of engineers.

It is important to continue to examine the experiences of women in engineering in a forum in which they can speak freely. In future work, we will carry out a similar text mining study on other social media platforms such as Twitter or Facebook. Furthermore, comments on platforms such as Glassdoor or LinkedIn can provide unique insight into specific employers and industries. Finally, as we mentioned earlier, comments on social media platforms have a tendency to be negative. An interesting direction for future work is to analyze positive discussions that specifically point out programs and initiatives to improve women's satisfaction and persistence in STEM.

REFERENCES

- [1] Catherine T Amelink and Elizabeth G Creamer. 2010. 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 (2010), 81–92.
- [2] Cornelia M Ashby. 2006. 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).
- [3] Christos Baziotis, Nikos Pelekis, and Christos Doukeridis. 2017. Datastories at semeval-2017 task 4: Deep lstm with attention for message-level and topic-based sentiment analysis. In *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*. 747–754.
- [4] Rae Lesser Blumberg. 2007. *Gender bias in textbooks: A hidden obstacle on the road to gender equality in education*. Unesco Paris.
- [5] Stephen G Brush. 2016. Women in Science and Engineering. *American Scientist* 79, 5 (Oct. 2016), 404–419.
- [6] Michelle J Budig, Paula England, et al. 2001. The wage penalty for motherhood. *American sociological review* 66, 2 (2001), 204–225.
- [7] C. Beil K. Dam N. Yassinskaya M. Shaw C. Mavriplis, R. Heller and C. Sorensen. 2010. Mind the Gap: Women in STEM Career Breaks. *Journal of Technology Management and Innovation* 5, 1 (2010).
- [8] Engineers Canada. [n.d.]. *Retention | Engineers Canada*. <https://engineerscanada.ca/diversity/women-in-engineering/retention>
- [9] Molly Carnes, Patricia G Devine, Carol Isaac, Linda Baier Manwell, Cecelia E Ford, Angela Byars-Winston, Eve Fine, and Jennifer Sheridan. 2012. Promoting institutional change through bias literacy. *Journal of Diversity in Higher Education* 5, 2 (2012), 63.
- [10] Sapna Cheryan and Victoria C Plaut. 2010. Explaining underrepresentation: A theory of precluded interest. *Sex roles* 63, 7-8 (2010), 475–488.
- [11] Shivangi Chopra, Hannah Gautreau, Abeer Khan, Melicaalsadat Mirsafian, and Lukasz Golab. 2018. Gender Differences in Undergraduate Engineering Applicants: A Text Mining Approach. In *Proceedings of the 11th International Conference on Educational Data Mining, EDM 2018, Buffalo, NY, USA, July 15-18, 2018*. 44–54.

- [12] Shivangi Chopra, Abeer Khan, Melicaalsadat Mirsafian, and Lukasz Golab. 2019. Gender Differences in Science and Engineering: A Data Mining Approach. In *Proceedings of the Workshops of the EDBT/ICDT 2019 Joint Conference, EDBT/ICDT 2019, Lisbon, Portugal, March 26, 2019*. http://ceur-ws.org/Vol-2322/DARLIAP_7.pdf
- [13] Shivangi Chopra, Abeer Khan, Melicaalsadat Mirsafian, and Lukasz Golab. 2019. Gender Differences in Work-Integrated Learning Assessments. In *Proceedings of the 12th International Conference on Educational Data Mining, EDM 2019, Montréal, Canada, July 2-5, 2019*. <https://drive.google.com/file/d/1W78yWFprzoQmazdrO1JArD6zZ8FhJt3F>
- [14] Jennifer Crocker, Andrew Karpinski, Diane M Quinn, and Sara K Chase. 2003. 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 (2003), 507.
- [15] Mumun De Choudhury and Sushovan De. 2014. Mental health discourse on reddit: Self-disclosure, social support, and anonymity. *Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014*, 71–80.
- [16] Mumun De Choudhury, Shagun Jhaver, Benjamin Sugar, and Ingmar Weber. 2016. Social Media Participation in an Activist Movement for Racial Equality. *Proceedings of the 10th International AAAI Conference on Weblogs and Social Media 2016*, 92–101.
- [17] Eric D Deemer, Dustin B Thoman, Justin P Chase, and Jessi L Smith. 2014. Feeling the threat: Stereotype threat as a contextual barrier to women's science career choice intentions. *Journal of Career Development* 41, 2 (2014), 141–158.
- [18] Alessandra Jayne Dinin. 2017. An Exploration in Theory of the Storied Experiences of Women Earning Engineering Bachelor's Degrees at a Southern, Research, Predominately White Institution. *ProQuest LLC* (2017).
- [19] Kuheli Dutt, Danielle L Pfaff, Ariel F Bernstein, Joseph S Dillard, and Caryn J Block. 2016. Gender differences in recommendation letters for postdoctoral fellowships in geoscience. *Nature Geoscience* 9, 11 (2016), 805.
- [20] Kimberly A Eddleston, David C Baldrige, and John F Veiga. 2004. Toward modeling the predictors of managerial career success: does gender matter? *Journal of Managerial Psychology* 19, 4 (2004), 360–385.
- [21] Richard M Felder, Gary N Felder, Meredith Mauney, Charles E Hamrin Jr, and E Jacquelin Dietz. 1995. A longitudinal study of engineering student performance and retention. III. Gender differences in student performance and attitudes. *Journal of Engineering Education* 84, 2 (1995), 151–163.
- [22] Kimberley Ferriman, David Lubinski, and Camilla P Benbow. 2009. 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 (2009), 517.
- [23] Darcy Hango. 2013. Gender Differences in science, technology, engineering, mathematics, and computer science (STEM) programs at university. *Insights on Canadian Society* (12 2013).
- [24] Susan Harkness and Jane Waldfogel. 2003. The family gap in pay: Evidence from seven industrialized countries. In *Worker well-being and public policy*. Emerald Group Publishing Limited, 369–413.
- [25] S. A. Hewlett. [n.d.]. *Executive Women and the Myth of Having It All*. <https://hbr.org/2002/04/executive-women-and-the-myth-of-having-it-all>.
- [26] Christopher Homan, Ravdeep Johar, Tong Liu, Megan Lytle, Vincent Silenzio, and Cecilia Ovesdotter Alm. 2014. Toward macro-insights for suicide prevention: Analyzing fine-grained distress at scale. In *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, 107–117.
- [27] Jennifer Hunt. 2016. Why do women leave science and engineering? *ILR Review* 69, 1 (2016), 199–226.
- [28] Aditya Johri, Habib Karbasian, Aqdas Malik, Rajat Handa, and Hemant Purohit. 2018. How diverse users and activities trigger connective action via social media: lessons from the twitter hashtag campaign# ilooklikeanengineer. *arXiv preprint arXiv:1804.09226* (2018).
- [29] Antti Kauhanen and Sami Napari. 2011. *Gender differences in careers*. Technical Report. ETLA discussion paper.
- [30] H.Liu L. Manikonda, G. Beigi and S.Kambhampati. 2018. *Twitter for Sparking a Movement, Reddit for Sharing the Moment: metoo through the Lens of Social Media*. Technical Report. Arizona State University.
- [31] Michael Thaul Lehrman, Cecilia Ovesdotter Alm, and Rubén A Proano. 2012. Detecting distressed and non-distressed affect states in short forum texts. In *Proceedings of the Second Workshop on Language in Social Media*. Association for Computational Linguistics, 9–18.
- [32] Judith Lorber, Susan A Farrell, et al. 1991. *The social construction of gender*. Sage Newbury Park, CA.
- [33] L. Subich M. Cadaret, P. Hartung and I. Weigold. 2017. Stereotype threat as a barrier to women entering engineering careers. *Journal of Vocational Behavior* 99, 1 (Oct. 2017), 40–51.
- [34] David MacPhee, Samantha Farro, and Silvia Sara Canetto. 2013. Academic self-efficacy and performance of underrepresented STEM majors: Gender, ethnic, and social class patterns. *Analyses of Social Issues and Public Policy* 13, 1 (2013), 347–369.
- [35] Katherine L Milkman, Modupe Akinola, and Dolly Chugh. 2015. 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 (2015), 1678.
- [36] K. Cappaert W. Chang N. Fouad, R. Singh and M. Wan. 2016. Comparison of women engineers who persist in or depart from engineering. *Journal of Vocational Behavior* 92, 1 (Oct. 2016), 79–93.
- [37] Rebecca Northfield. 2018. Female engineers# MeToo? *Engineering & Technology* 13, 6 (2018), 32–36.
- [38] Derek O'callaghan, Derek Greene, Joe Carthy, and Pádraig Cunningham. 2015. An analysis of the coherence of descriptors in topic modeling. *Expert Systems with Applications* 42, 13 (2015), 5645–5657.
- [39] Frank Pajares. 2005. Gender differences in mathematics self-efficacy beliefs. *Gender differences in mathematics: An integrative psychological approach* (2005), 294–315.
- [40] V Paul Pauca, Farial Shahnaz, Michael W Berry, and Robert J Plemmons. 2004. Text mining using non-negative matrix factorizations. In *Proceedings of the 2004 SIAM International Conference on Data Mining*. SIAM, 452–456.
- [41] Umashanthi Pavalanathan and Mumun De Choudhury. 2015. Identity management and mental health discourse in social media. In *Proceedings of the 24th International Conference on World Wide Web*. ACM, 315–321.
- [42] Abigail Powell, Barbara Bagilhole, and Andrew Dainty. 2009. How women engineers do and undo gender: Consequences for gender equality. *Gender, work & organization* 16, 4 (2009), 411–428.
- [43] B. Chakalov S. Modrek. 2019. The MeToo Movement in the United States: Text Analysis of Early Twitter Conversations. *Journal of Medical Internet Research* 21, 9 (Sept. 2019). <https://doi.org/10.2196/13837>
- [44] Linda J Sax. 1994. Mathematical self-concept: How college reinforces the gender gap. *Research in Higher Education* 35, 2 (1994), 141–166.
- [45] Linda J Sax. 1994. Predicting gender and major-field differences in mathematical self-concept during college. *Journal of Women and Minorities in Science and Engineering* 1, 4 (1994).
- [46] Nicolas Schradang, Cecilia Ovesdotter Alm, Raymond Ptucha, and Christopher Homan. 2015. # whyistayed.# whyleft: Microblogging to make sense of domestic abuse. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 1281–1286.
- [47] Nicolas Schradang, Cecilia Ovesdotter Alm, Ray Ptucha, and Christopher Homan. 2015. An analysis of domestic abuse discourse on Reddit. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. 2577–2583.
- [48] Carroll Seron, Susan S Silbey, Erin Cech, and Brian Rubineau. 2016. Persistence is cultural: Professional socialization and the reproduction of sex segregation. *Work and Occupations* 43, 2 (2016), 178–214.
- [49] Adrienne Y. Smith. 2012. *They chose to major in engineering: A study of why women enter and persist in undergraduate engineering programs*. Ph.D. Dissertation. University of Massachusetts - Amherst.
- [50] Jennifer Steele, Jacquelyn B James, and Rosalind Chait Barnett. 2002. Learning in a Man's world: Examining the perceptions of undergraduate women in Male-Dominated academic areas. *Psychology of women quarterly* 26, 1 (2002), 46–50.
- [51] IPEDS Completion Survey. [n.d.]. American Physical Society. <https://www.aps.org/programs/education/statistics/womenmajors.cfm>. Accessed: 20th August, 2018.
- [52] Divy Thakkar, Nithya Sambasivan, Purva Kulkarni, Pratap Kalenahalli Sudarshan, and Kentaro Toyama. 2018. 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*. ACM, 352.
- [53] Joachim Tiedemann. 2000. 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 (2000), 144.
- [54] Joseph B Walther. 1996. Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication research* 23, 1 (1996), 3–43.
- [55] Ming-Te Wang and Jessica L. Degol. 2017. Gender Gap in Science, Technology, Engineering, and Mathematics (STEM): Current Knowledge, Implications for Practice, Policy, and Future Directions. *Educational Psychology Review* 29, 1 (01 Mar 2017), 119–140. <https://doi.org/10.1007/s10648-015-9355-x>
- [56] Joan Williams, Katherine Phillips, and Erika Hall. 2014. *Double Jeopardy? Gender Bias Against Women of Color in Science*. Published online at www.worklifelaw.org. <https://doi.org/10.13140/2.1.1763.8723>
- [57] Shelly L Wismath and Maggie Zhong. 2014. 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).
- [58] Jun-Ming Xu, Benjamin Burchfiel, Xiaojin Zhu, and Amy Bellmore. 2013. An examination of regret in bullying tweets. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 697–702.