Engineering Educators Bringing the World Together
2025 ASEE Annual Conference & Exposition
Palais des congrès de Montréal, Montréal, QC • June 22–25, 2025
Paper ID #47967

Hidden Trends in Data on Women in STEM

Dr. Laura Bottomley, North Carolina State University at Raleigh

Dr. Laura Bottomley is the Director of Engineering Education and Senior Advisor to Engineering Academic Affairs at NC State University. She is also an Adjunct Professor for Electrical and Computer Engineering at NC State. She has been working in the field of engineering education for more than 30 years, having taught every grade level from kindergarten to engineering graduate school. She is a Fellow of the IEEE and ASEE and has been recognized with the PAESMEM award.

Dr. Leah Granger, North Carolina State University at Raleigh

Leah Granger is a postdoctoral researcher for Engineering Education and a course instructor for the Chemical and Biomolecular Engineering Department at North Carolina State University.

Hidden Trends in Data on Women in STEM

Introduction

The use of data to monitor progress in the recruitment and retention of underrepresented populations in STEM encourages careful consideration of the manner in which data are grouped in the analysis. Trends present in the overall population of study – for example, college students enrolled in a STEM program – may not be an accurate reflection of trends in specific subpopulations. Numerically, majority populations have a larger influence on the observed patterns of the overall dataset, necessitating analysis of subpopulations in order to identify group-specific trends. Example subpopulations include the intersection of a given gender and race or may be defined by specific fields within STEM. For example, the trends observed in biological sciences may not reflect those of physical sciences, especially regarding gender disparities. The present paper includes two case studies, one using university data and one using nationwide data, to demonstrate discrepancies in STEM participation trends in postsecondary education based on the specificity of the group. These are followed by a discussion on the challenges associated with managing data through the STEM pipeline as the programs defined as STEM at colleges and universities do not always align with occupations classified as STEM in national databases. These inconsistencies add difficulty in tracking the retention of STEM graduates in STEM careers, and consequently, hinders studies on the challenges and barriers faced by underrepresented populations throughout their educational and professional careers.

Examination of the literature

The literature reveals some scattered efforts to look at data through different lenses. Byrd, et al. [1] in 2013 looked at national data in the Integrated Postsecondary Education Data System (IPEDS) dataset. They argued that forcing non-US students into a group category ignored their race/ethnicity, which might significantly bias the data set. They suggested that data should be collected looking at the intersection of race/ethnicity and nationality/citizenship.

In 2017, Ma and Liu [2] used the National Education Longitudinal Studies (1988:2000) (NELS) database from the National Center for Education Statistics to look at degree attainment in STEM fields. They looked at the data from an intersectional perspective, using male/female gender and non-Hispanic White, non-Hispanic Black, Hispanic and Asian race/ethnic categories.

Ro and Loya [3] looked at how students rate their attainment of learning outcomes by gender combined with race using the Prototype to Production: Conditions and Processes for Educating the Engineer of 2020 (P2P) dataset, while Lord et al. [4] used the Multiple-Institution Database for Investigating Engineering Longitudinal Development (MIDFIELD) dataset to examine persistence in engineering by race/ethnicity and gender. A few more such studies exist in the literature [5, for example], using different data sets and looking at different subsets of the STEM population.

This study carries forth the idea that differences in outcomes may exist when data are examined through an intersectional lens, and takes the examination further by asking the question, what are we missing by NOT using this approach.

University-scale case study:
Engineering enrollment trends based on gender, race, and gender + race

Enrollment in the College of Engineering (COE) at North Carolina State University has grown in recent years, as illustrated in Figure 1. However, when considering data from 2016 through 2024, enrollment experienced two noticeable stalls in growth: the first beginning in 2017 (before returning to its high growth rate between 2019 and 2020) and a second stall beginning in 2020 (with recovery beginning in 2022). While the cause of each of these instances of growth stall may be attributed to intentional University decisions and/or external factors, the following discussion does not seek to explain the causes. Rather, the discussion looks at how enrollment of specific subpopulations contributed to these observed growth stalls, serving as an exploratory case study as to how data analysis for subpopulations can provide important insights into otherwise hidden trends that are specific to those groups.

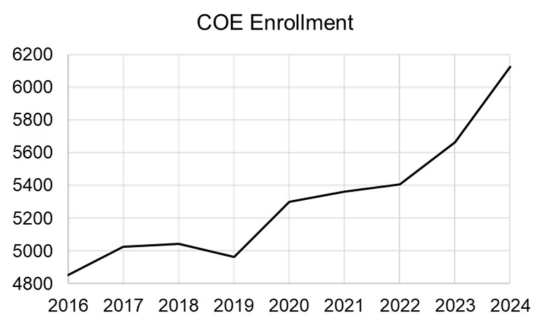

Figure 1: Total enrollment in the College of Engineering. Note that these enrollment numbers only include students who have been accepted into their engineering major and exclude undecided and intended majors.

Comparing enrollment by gender indicates the 2017 stall was less severe for female enrollment, but its 2020 stall was more prolonged, as shown in Figure 2. In the 2017 growth stall, female enrollment growth did not experience an observable change until 2018 and recovered with male enrollment between 2019 and 2020. This indicates the contributing factors of this event affected males more than females. Male and female enrollment showed similar growth decline beginning in 2020, but female enrollment recovered a year later than male enrollment. The delayed recovery added to the ongoing challenge of reducing gender disparity. Although the fraction of females enrolled is at its historic maximum – 28.3% of COE students identify as female – overall male enrollment is currently increasing at a faster rate than female enrollment, so closing the gender gap remains a challenge. As demonstrated in Figure 3, within racial groups the gender gap is most pronounced in white and Hispanic enrollments, with female enrollment representing only 26.5% of white enrollment and 24.2% of Hispanic enrollment in 2024. Female enrollment exceeds 35% in African American and Asian enrollment, with the gender gap in Asian enrollment consistently shrinking since 2021.

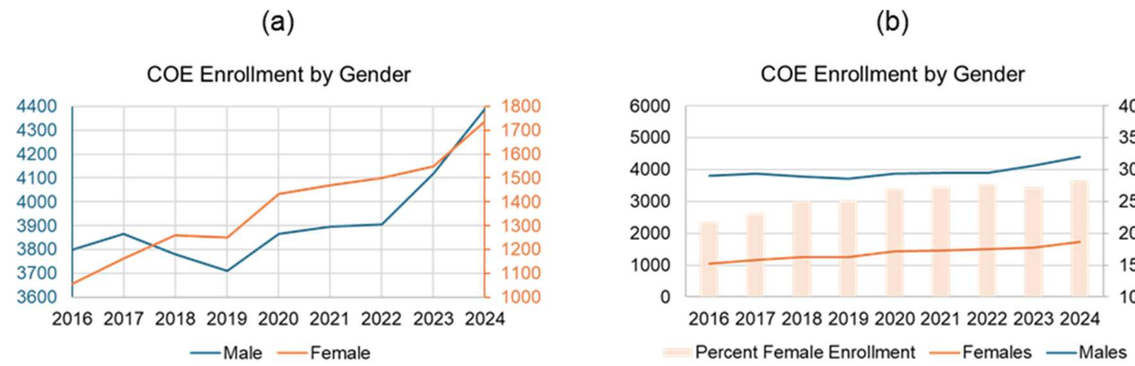

Figure 2: (a) COE enrollment by gender where male enrollment is plotted on the left y-axis and female enrollment is plotted on the right y-axis. Note that both y-axes have a range of 800 students to allow a comparison of slopes. (b) COE enrollment by gender plotted on the left y-axis, and the percent of total enrollment attributed to female-identifying students plotted on the right y-axis.

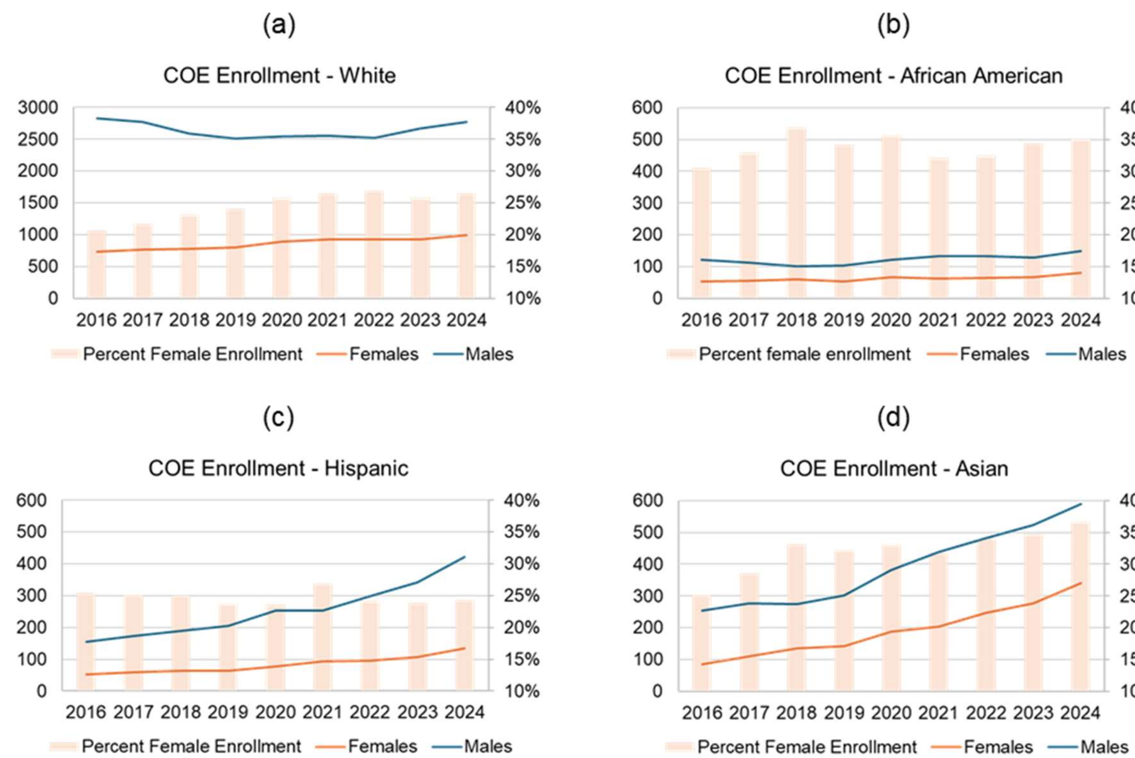

Figure 3: COE enrollment based on gender for (a) white, (b) African American, (c) Hispanic, and (d) Asian enrollments, where enrollments based on gender are plotted on the left y-axis, and the percent of enrollment attributed to female-identifying students within the given racial group is plotted on the right y-axis. Note that the y-axes for plots (b), (c), and (d) are equivalent to allow a comparison of enrollment and rate of growth.

If enrollment data is organized by race, the 2017 and 2020 growth stalls are predominantly observed in enrollment of students identifying as white, as illustrated in Figure 4. Note that only white, African American, Hispanic, and Asian enrollments are included in this analysis due to the small enrollment numbers for other groups that make comparison and trend analysis more challenging.

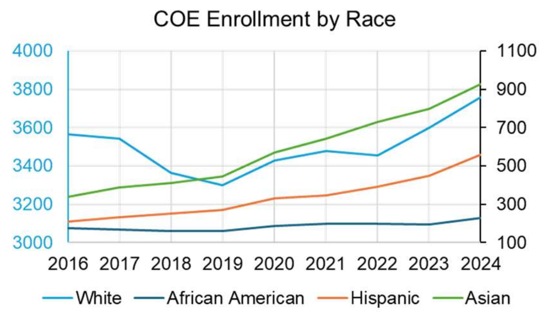

Figure 4: COE enrollment by race where white enrollment is plotted on the left y-axis and African American, Hispanic, and Asian enrollments are plotted on the right y-axis. Note that both y-axes have a range of 1000 students to allow a comparison of slopes.

Analysis of enrollment by both race and gender, as done in Figure 5, indicates that white males experience the most significant changes in enrollment. White females also experience significant change relative to other racial groups, but the effect is less pronounced as white female enrollment stagnates while white male enrollment decreases. Further, white female enrollment exhibits the two distinct 2017 and 2020 phases of growth stall, but male enrollment appears to have limited recovery between the two time periods.

Trends observed in other racial groups contribute to overall trend behaviors. Asian male enrollment growth stalls in 2017 but is mostly unaffected in 2020, whereas Hispanic male enrollment experienced the reverse, with minimal change in 2017 but a noticeable stall in 2020. Changes in female enrollments appear less drastic compared to those of male enrollments. Each subpopulation experiences two periods of minor growth stagnation, although Asian female enrollment stagnated in 2018 rather than 2017, and Hispanic female enrollment stagnated in 2021 rather than 2020 – yearlong delays that could possibly indicate there are multiple contributing factors. African American males, African American females, and white females did not return to pre-2020 growth rates until 2023, one to two years later than other groups. The overall trend demonstrated by Figure 1 and gender-based trends shown in Figure 2 are not accurate descriptions for all subpopulations, highlighting the need for deeper analysis.

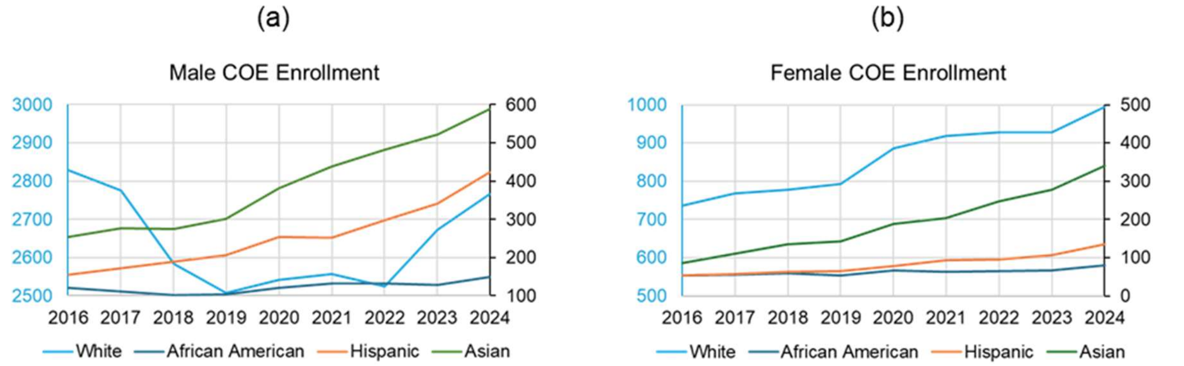

Figure 5: COE enrollment by race for (a) male-identifying students and (b) female-identifying students, where white enrollment is plotted on the left y-axis and African American, Hispanic, and Asian enrollments are plotted on the right y-axis. Note that all y-axes have a range of 500 students to allow a comparison of slopes.

This case study demonstrates the need to consider subpopulations – such as the intersection of a gender and race – in data analysis, but there is also a need to account for subfields within engineering to identify trends specific to certain fields. A balance must be found, however, between dataset size and subset specificity; analysis for specific subgroups can inform advocacy and programming efforts, but a large enough sample size is needed to identify significant trends.

Nationwide-scale case study:
S&E degrees awarded based on gender, race, and gender + race

The National Science Foundation report, *Diversity and STEM: Women, Minorities, and Persons with Disabilities 2023* [6], is a common source of information on representation in STEM fields. The report analyzes data from the National Center for Science and Engineering Statistics (NCSES). The data tables created for the report are available for download by gender, race/ethnicity, and citizenship status by each of the STEM categories. The STEM categories included in this data include science and engineering (S&E) and non-S&E fields.

The data are very complete, but are provided as an adjunct to the report. Both media and researchers may be inclined to look for statements in the report instead of the data. In the following sections are two statements from the report, followed by analyses from the data. The statements, while true, might lead to either a misunderstanding or a faulty conclusion about the state of representation in STEM fields.

The first statement is found in the Introduction/Overview of the report:

"Women earned approximately half [48%] of the S&E degrees at the associate's and bachelor's degree levels in 2020, which was similar to their share of the population ages 18 to 34 years."

Figure 6 was created from analysis of the report data and shows the representation of women in some sample disciplines within the S&E category. When considering all S&E degrees enumerated in the report, males and females are represented equally. It will not be a surprise to those working in these fields that this representation varies widely across engineering, physics and biology, which are not the most extreme examples.

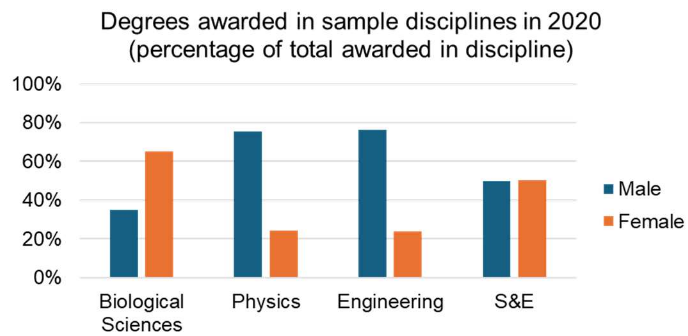

Figure 6: Degrees granted in 2020 in S&E subsets and broken out by gender. For reference, the total number of degrees granted in 2020 in biological sciences was 133,109, in physics was 7,714, in engineering was 131,062, and in S&E was 745,110.

There are short statements elsewhere in the report that reference this kind of variability, but they are separated from the statement quoted above. As an example, in the eighth section of the document, women are described as having received 66% of the bachelor's degrees awarded in social and behavioral sciences, 64% in agricultural and biological sciences, 26% of degrees in mathematics and computer sciences, and 24% in engineering. Even these statistics leave out the wide variability *within* these broad fields.

The data analyzed within the report contain some interesting information about the intersection between sex and race/ethnicity that is not covered in the report. The second example statement, referring to degrees awarded in 2020, examined here suggests the need to look more closely at these data:

"Underrepresented minorities collectively accounted for 37% of the college-age population in 2021 and 26% of S&E bachelor's…degree recipients."

The data show considerable variation from one category of underrepresented population to another. In Figure 7, below, degrees awarded to women in five categories of race/ethnicity are graphed over time from 2011 to 2020. Hispanic/Latina women more than doubled the number of degrees received over the ten years graphed. This line representing this category of student has a slope of almost 1% point per year, compared to the line representing students of more than one race, whose slope is almost .4% per year. The two lines representing native populations, by contrast, have negative slopes.

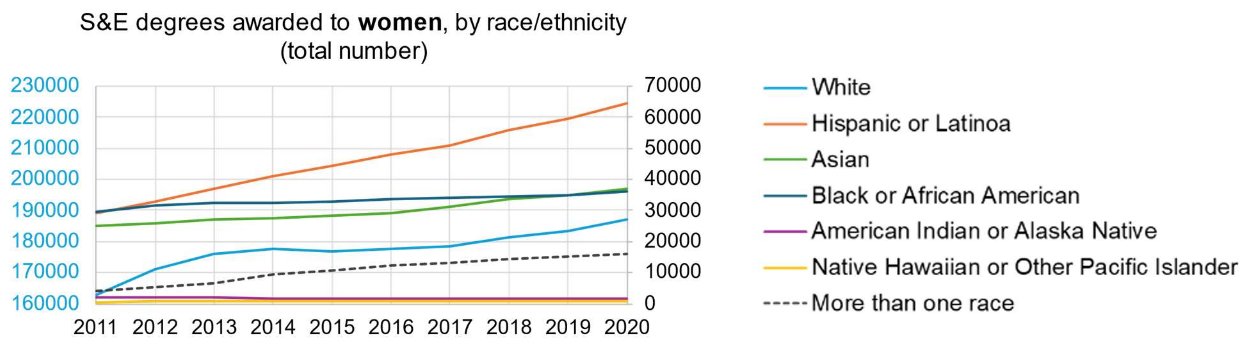

Figure 7: S&E degrees granted to women, plotted by race/ethnicity for years 2011-2020, where degrees awarded to white women is plotted on the left y-axis and degrees awarded to other races/ethnicities are plotted on the right y-axis. Note that both y-axes have a range of 70,000 degrees to allow a comparison of slopes.

Analyzing the data based on race/ethnicity reveals important information about which groups are holding steady and which are growing, or shrinking, as the total population grows. This has always been one of the difficulties in maintaining a diverse population of students. As the overall population grows, maintaining or increasing the proportion of women in science and engineering is challenging. Looking at the data on women divided by race/ethnicity tells a deeper story that might help us identify root causes for the lack of representation.

Figures 8 and 9 include the data from Figure 6 above as a subset. These plots show how the proportion of degrees awarded to white women and men has changed to accommodate the percentage growth of other racial/ethnic groups. Note that this does not mean that the *number* of degrees awarded to white men and women has decreased; they have increased. The graphs show numbers represented as a percentage of the total population of each racial/ethnic group.

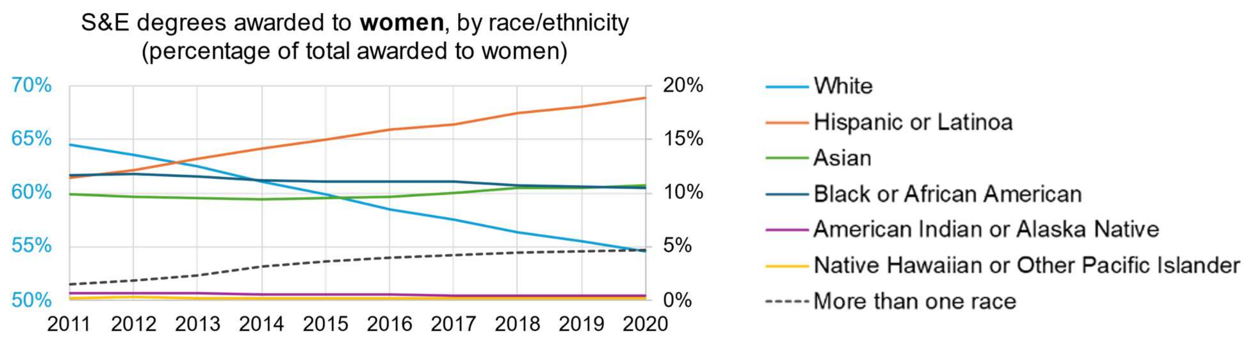

Figure 8: Percentage of S&E Degrees awarded to women of various race/ethnicity as a percentage of total S&E degrees awarded to women, where the percentage of degrees for white women is plotted on the left y-axis and the percentages of degrees for other races/ethnicities are plotted on the right y-axis. Note that both y-axes have a range of 20% of the total to allow a comparison of slopes.

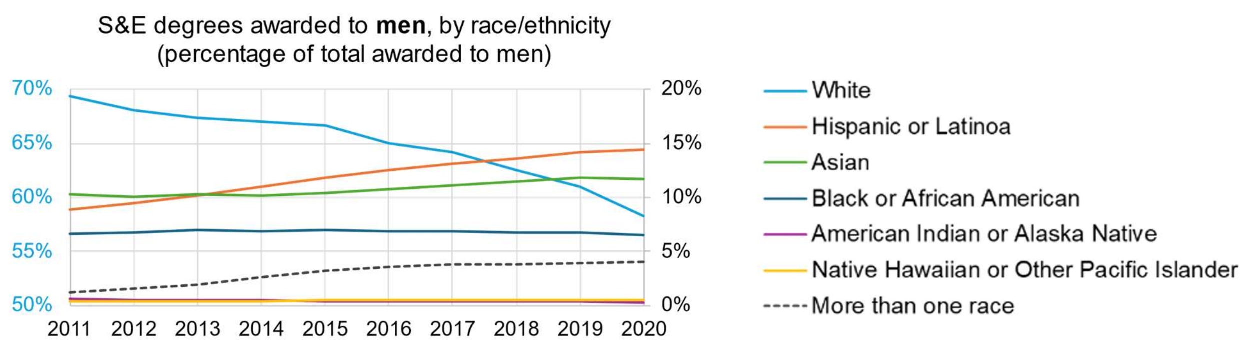

Figure 9: Percentage of S&E Degrees awarded to men of various race/ethnicity as a percentage of total S&E degrees awarded to men, where the percentage of degrees for white men is plotted on the left y-axis and the percentages of degrees for other races/ethnicities are plotted on the right y-axis. Note that both y-axes have a range of 20% of the total to allow a comparison of slopes.

Examining these two graphs side by side shows that the changes in percentage of degrees awarded by group looks different by gender. Asian women are not showing gains as high as Asian men. Hispanic/Latino men are not showing gains as high as Hispanic/Latina women. Overall, women who get degrees in science and engineering are more ethnically/racially diverse than men. These are important nuances.

The differences by sex are emphasized for four of the racial/ethnic groups in Figure 10, showing the degree to which female Hispanic/Latinas are exceeding their male counterparts as in representation. Similarly, female Black/African American students are exceeding their male counterparts. White women and Asian women trail their male counterparts.

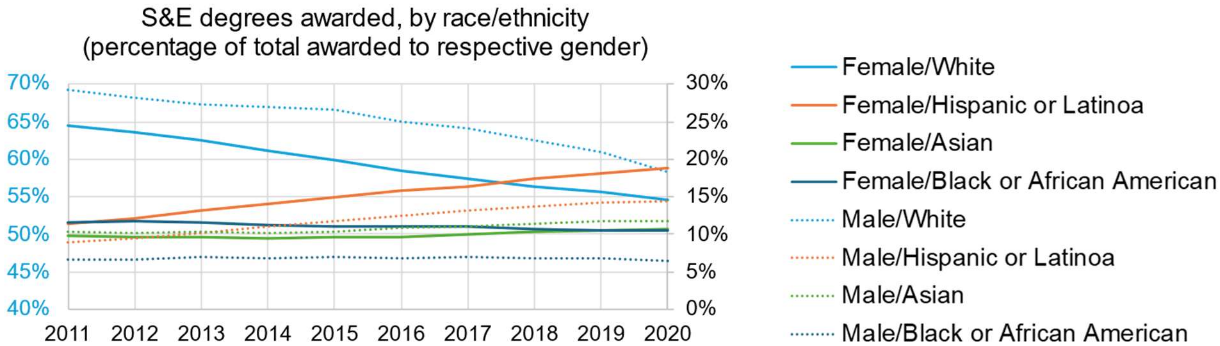

Figure 10: Science and engineering degrees awarded by intersectional race/ethnicity and gender, as a percentage of degrees awarded to their respective gender, where the percentage of degrees for white men and women is plotted on the left y-axis and the percentages of degrees for other races/ethnicities are plotted on the right y-axis. Note that both y-axes have a range of 30% of the total to allow a comparison of slopes.

Analyzing the definitions of STEM in the ecosystem

An interesting complication in understanding journeys in the STEM ecosystem is differences in the definitions of "STEM" at various levels. For the purposes of this paper, we will look at two levels: undergraduate degree granting and employment. A logical place to look for progress for women and other minoritized groups is the transition from graduation with a BS degree in a STEM field and employment.

The definition of STEM varies by data source. Notably, the definition of what constitutes STEM degrees is actually different from the definition used to talk about STEM occupations. So, for example, the findings in *Science and Engineering Indicators 2024* [7] on the demographic composition of the STEM workforce and those on science and engineering higher education uses two different definitions of science, engineering, and STEM.

While there is no universal definition, a common method for the categorization of fields as STEM used in higher education comes from the Integrated Postsecondary Education Database System (IPEDS) [8] and Classification of Instructional Programs (CIP) codes [9]. Broadly speaking, for discussion of degrees in higher education, the classifications of science and engineering (S&E), non-S&E, and S&E-related fields are used. More discussion of what exactly constitutes these fields is available in the "Higher Education in Science and Engineering" section of *Science and Engineering Indicators 2024* [7]. These definitions have also changed over time, so comparing data between reporting years may not be accurate. At a high level, science and engineering fields include CIP codes:

- **CIP 01:** Agricultural sciences and natural resources
- **CIP 11:** Computer and Information Sciences and Support Services
- **CIP 14:** Engineering
- **CIP 15:** Engineering Technologies/Technicians
- **CIP 26:** Biological and Biomedical Sciences
- **CIP 27:** Mathematics and Statistics
- **some of CIP 30:** Multidisciplinary and interdisciplinary sciences
- **CIP 40:** Physical Sciences (e.g., Physics, Chemistry, Earth Sciences)
- **CIP 41:** Science Technologies/Technicians
- **CIP 42**: Psychology
- **CIP 45**: Social Sciences

This level of detail is very important when comparing across levels of the STEM ecosystem. Most academics are probably familiar with this classification system, even if they have not dealt with CIP codes specifically. Interestingly, when looking closely at the way this system is applied in Table 2-2 of the National Science Foundation's (NSF) report on *Diversity and STEM: Women, Minorities, and Persons with Disabilities* [6], some degrees that are explicitly excluded from S&E degrees are “Science and engineering technologies,” including “Engineering technologies,” “Health technologies,” “Science technologies,” and “Other science and engineering technologies.” Some areas that are included in S&E degrees, that may be surprising to some, include “Area and ethnic studies,” “History of science,” “Linguistics,” and “Political science and public administration.”

The definition of STEM occupations differs from the definitions of science and engineering fields used to describe higher education. US Census Bureau data uses OOC codes [10] to define STEM-based occupations to S&E occupations, S&E-related occupations, and middle-skill occupations. These codes are quite detailed, and a complete discussion of them exceeds the capacity of this paper, but a few examples are illustrative. It is important to note that OOC codes are NOT correlated with CIP codes. The table below contains some examples of occupations that fit into each of the three categories included in this definition of STEM.

Table 1: A sample of OOC codes categorizing S&E, S&E-related, and middle-skill occupations.

S&E	
1005	Computer and information research scientists
1050	Computer support specialists
1210	Mathematicians
1320	Aerospace engineers
1440	Marine engineers and naval architects
1600	Agricultural and food scientists
1660	Life scientists, all others
1760	Physical scientists, all other
1815	Survey researchers
1820	Psychologists
1830	Sociologists
1840	Urban and regional planners
S&E related	
360	Natural sci managers
1010	Computer programmers
1200	Actuaries
1300	Architects, except Naval
1550	Engr tech, except drafters
1900	Agri and food science technicians
1950	Social science research assistants
1965	Miscellaneous life, physical, and social science technicians
3010	Dentists
3060	Physicians and surgeons
3160	Physical therapists
3255	Registered nurses
Middle skilled	
140	Industrial production managers
205	Farmers, ranchers, and other agricultural managers
2840	Technical writers
2900	Broadcast and sound engineering technicians and radio operators, and media and communication equipment workers
3740	Firefighters
4000	Chefs and head cooks
6210	Boilermakers
6330	Drywall installers, ceiling tile installers, and tapers
6355	Electricians
6360	Glaziers
6820	Earth drillers, except oil and gas
7810	Butchers and other meat, poultry, and fish processing workers

The discrepancy in STEM definitions for higher education and for occupations adds challenges when identifying retention trends. Further, the categorizations of specific fields and occupations is not always intuitive – excluding some commonly considered STEM fields while including some non-traditional fields – and can obscure trends, especially when looking for data concerning underrepresented groups.

One of the biggest implications of the above differing definitions is how it complicates efforts to trace population changes through the STEM ecosystem. Numerous publications quote the number of STEM degrees awarded to women and then describe how the percentages of women working in STEM fields are much lower. The differing definitions of STEM at each of those locations in the ecosystem mean that the statement cannot be taken at face value.

Conclusions

Efforts to increase underrepresented populations pursuing degrees and careers in STEM fields rely on data to identify trends in recruitment and retention; however, when data is analyzed in large datasets, trends present in these smaller, underrepresented populations can be obscured by those of the majority population. Data that groups several fields together can also bury significant trends present only in specific fields. Careful consideration of the appropriate data subsets is necessary to uncover otherwise hidden patterns. Further, the STEM ecosystem, with its pathways from postsecondary education to the workforce, is challenging to follow as there is no universal definition of STEM, and the classifications of STEM programs in higher education do not align with those used to define STEM occupations. Educators and administrators must be careful when comparing data sets to uncover a more complete story and better address the challenges and barriers facing underrepresented students.

Acknowledgements

The authors would like to thank the North Carolina State University College of Engineering Factbook for providing COE enrollment data for 2016 – 2024.

References

[1] Byrd, W. C., Dika, S. L., & Ramlal, L. T. (2013). Who's in STEM? An Exploration of Race, Ethnicity, and Citizenship Reporting in a Federal Education Dataset. Equity & Excellence in Education, 46(4), 484–501. https://doi.org/10.1080/10665684.2013.838485

[2] Ma, Y., & Liu, Y. (2017). Entry and Degree Attainment in STEM: The Intersection of Gender and Race/Ethnicity. Social Sciences, 6(3), 89. https://doi.org/10.3390/socsci6030089

[3] Ro, H.K., & Loya, K.I. (2015). The Effect of Gender and Race Intersectionality on Student Learning Outcomes In Engineering. The Review of Higher Education 38(3), 359-396. https://dx.doi.org/10.1353/rhe.2015.0014.

[4] Lord, S. M., Camacho, M. M., Layton, R. A., Long, R. A., Ohland, M. W., & Washburn, M. H. (2009). Who's persisting in engineering? A comparative analysis of female and male Asian, Black, Hispanic, Native American, and White students. Journal of Women and Minorities in Science and Engineering, 15(2), 167–190.

[5] Lord, S. M., Layton, R. A., & Ohland, M. W. (2011). Trajectories of electrical engineering and computer engineering students by race and gender. IEEE Transactions on Education, 54(4), 610–618. doi: 0018-9359

[6] National Center for Science and Engineering Statistics (NCSES). 2023. *Diversity and STEM: Women, Minorities, and Persons with Disabilities 2023*. Special Report NSF 23-315. Alexandria, VA: National Science Foundation. Available at https://ncses.nsf.gov/wmpd.

[7] National Science Board, National Science Foundation. 2019. Science and Engineering Indicators 2020: Science and Engineering Labor Force. *Science and Engineering Indicators 2020*. NSB-2019-8. Alexandria, VA. Available at https://ncses.nsf.gov/pubs/nsb20198/.

[8] National Center for Education Statistics (NCES). Integrated Postsecondary Education Data System. Available at https://nces.ed.gov/ipeds/.

[9] National Center for Education Statistics (NCES). The Classification of Instructional Programs. Available at https://nces.ed.gov/ipeds/cipcode/.

[10] US Census Bureau, Industry and Occupation Code Lists & Crosswalks. Available at https://www.census.gov/topics/employment/industry-occupation/guidance/code-lists.html.