Measuring and Analyzing the Gender Gap in Science through the Global Survey of Scientists

Measuring and analyzing the gender gap in science through the global survey of scientists

Rachel Ivie[1], Susan White[2]

1 – Senior Director of Education and Research, American Institute of Physics, College Park, MD, US
2 – Interim Director, Statistical Research Center, American Institute of Physics, College Park, MD, US

1 Methodology

The 2018 Global Survey of Mathematical, Natural, and Computing Scientists seeks to develop a broader picture of the status of mathematicians and scientists across the world. The survey instrument was developed by the Gender Gap in Science Project in collaboration with the American Institute of Physics (AIP). Various questions address specific developmental periods in the education and careers of scientists. The UNESCO Institute for Statistics reports that women constitute less than 30% of scientific researchers worldwide, despite women accounting for 50% of the global population. To understand representation of women in STEM fields, it is important to look at three life phases: 1) childhood, 2) early adulthood, and 3) professional life. The Global Survey takes these phases into account by assessing how scientists perceive their early years, university studies, doctoral studies, and careers.

This document explores gender differences by scientific discipline, regions, and level of human development. The disciplines included are Astronomy, Biology, Chemistry, Computer Science, Mathematics, Mathematics – Applied, and Physics. History and Philosophy of Science, although not a mathematical or natural science, is included in some

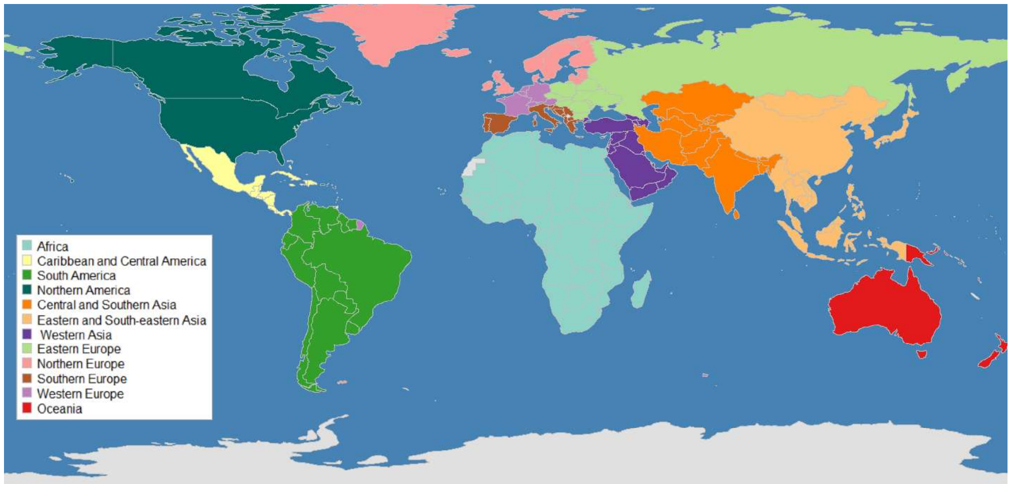


Figure 1: Geographic regions used in the analyses

of the analysis because the International Union of History and Philosophy of Science and Technology was a partnering organization. Low response rates prevented including this discipline in all analyses.

We also examined gender differences within regions to look at trends based on geography, not subject matter. We analyzed the data across twelve regions: Africa, Northern America, Caribbean and Central America, South America, West Asia, Central and Southern Asia, Eastern and South-eastern Asia, Northern Europe, Western Europe, Eastern Europe, Southern Europe, and Oceania (see Figure 1, p. 40).

Finally, we examined gender differences within and across human development levels. Development levels were based on the Human Development Index (HDI) created by the United Nations Development Programme. The index is a composite score that considers health, education, and standard of living in a country.[1] Based on this list, we divided the countries into two categories of development: more developed and less developed. Taiwan is the only country in the survey not included in the UN list. Based on available data on Taiwan's health, education and standard of living, Taiwan was included in the more developed category.

[1] United Nations Development Programme, *United Nations Human Development Index*, http://hdr.undp.org/en/content/human-development-index-hdi.

1.1 Survey Design

Goals, priorities, and research questions for the survey were discussed at the launch meeting of the project in June, 2017 in Paris. The overall research questions are: to better understand scientists' development of interest in science, experiences in education and careers, work-life balance, family support, demographics, access to resources needed to conduct science, and opportunities to contribute to the scientific enterprise. The research questions also included the need to make comparisons across regions, disciplines, and level of human development. Using these goals, AIP drafted a questionnaire based largely on the previously used Global Survey of Physicists[2] and the UNESCO SAGA questionnaire [6]. AIP reviewed each drafted question to ensure that it 1) met the goals agreed upon at the Paris meeting, 2) contributed significantly to findings, and 3) answered the research questions.

AIP created a draft survey to present at regional meetings held in Colombia, Taiwan, and South Africa. Representatives in these meetings reviewed and provided feedback on the initial draft, including 1) reviewing specific questions to collect feedback on regional implications of wording and topics, 2) input on the full survey instrument with special consideration to ensure that the questions work for the region and for all disciplines, and 3) an outline of the distribution plan. Feedback and comments from each of the meetings were recorded, compiled, and reviewed to edit and update the questionnaire. The edited questionnaire was presented back to the project's executive committee for final approval. After final approval the questionnaire was translated and placed online for distribution.

1.2 Questionnaire Translation

After the executive committee approved the final version of the questionnaire, it was sent to a translation company that specializes in translating questionnaires. The professional translation company utilized a three-step process with native speakers of each language. It involves initial translation, followed by a second editing by a second linguist, and finally proofreading by a third linguist. The professional translations were reviewed by project participants to ensure that the translations met the needs of scientists. The reviews by project participants were used to inform the professional translators and refine the transla-

[2] For more on the 2009 IUPAP Global Survey of Physicists, see [3, 4].

tions for the project audience. The questionnaire was translated into Spanish, Russian, French, Chinese, Japanese, and Arabic. The languages were chosen based on input from the participants at the workshops, recommendations from the UN, and on the languages used for IUPAP's Global Survey of Physicists in 2009.

1.3 Sampling

We collected data using a snowball sampling method and contact databases from partnering organizations to reach students and professional scientists across the globe. Since there is not a single network or resource available to contact all students and professional scientists globally, we used snowball sampling to take advantage of as many personal networks as possible.

Snowball sampling is a non-probability method for data collection and does not result in a statistically representative sample. Because of this, there are important limitations of analysis and interpretation for the data collected by the survey. The findings presented in this report should not be assumed to be representative of the intended population as a whole. Therefore, the findings below only indicate trends for the individuals who responded to the survey, not the overall population.

1.4 Data Processing

Primary data collection ended on December 31, 2018. AIP cleaned and prepared the data by finding and solving inconsistencies, creating new variables for analysis, and labeling variables and responses, thereby increasing the quality of the dataset. The process consisted of three steps, (1) data cleaning, (2) translation of open-ended responses, and (3) product preparation. Data cleaning involved processes designed to increase data quality and ensure the accuracy of analyses. We examined each variable individually and through valid groupings of variables to ensure consistency and believability of responses. Skip patterns were logged to allow for a review of the total number of respondents answering each question. Respondents with potential issues were also investigated. For translation, the open-ended responses for each question were linked to the unique identifier for each respondent in order to be able to pair the translated text with a respondent's other answers.

There were approximately 250,000 individual words that required translation. Translation was conducted by a professional translation service.

1.5 Analysis Methods

We initially conducted bivariate analyses to provide a simple view of the data. However, bivariate analyses can be confounded by intervening factors, leading to incorrect interpretations.[3] For the final report, we focus on multivariate analyses. The multivariate analyses allow the inclusion of potential confounding factors, such as employment sector, discipline, geographic region, age and more in the analysis. We are still able to test for statistically significant differences in the responses of men and women after accounting for potential confounding factors. In this final report, we focus solely on the multivariate analyses to avoid the possibility of incorrect interpretations resulting from confounding. We have included the graphs of bivariate results where they are consistent with the multivariate results. However, in almost every case, the bivariate analysis understates the relative ratio of men's and women's responses.

Logistic and Ordinal Logistic Regression

We use logistic and ordinal logistic regression analyses [2] to examine gender difference by discipline, by region, and by level of economic development.

In cases where the variable of interest is binary (as in yes/no), we use logistic regression. In this model, we are estimating the log-odds of the event that the variable of interest = 1 (and not 0). We assume a linear relationship of the form:

$$l = \log_b \frac{p}{1-p} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + e,$$

where l is the log-odds, b is the base of the logarithm, β_i are the parameters of the model on the independent variables x_i, e represents random error, and p is the probability that the value of interest is 1 (and not 0).

[3] For example, we see this confounding effect in the analyses of doctoral program quality ranking. The bivariate analyses did not indicate any gender differences; however, the multivariate analysis reveals that men are likely to rate their doctoral program quality higher than women.

After the application of algebraic manipulation techniques, this resolves to our final model

$$p = \frac{1}{1 + b^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k)}}.$$

In the case of a binary dependent variable, we can use the output of the analysis to express the relative likelihood of one event over another. We do so in the tables.

In cases where the variable of interest has multiple options (for example, level of agreement), we use a modified version of this technique called ordinal logistic regression. In these models, the relative likelihood of an event depends on the number of steps of change in the response. For example, a movement from disagree to strongly agree would be three steps in the case of a five-point scale with a neutral point in the middle. Likewise, a change in the response from strongly disagree to agree would also be three steps. So, it is not as straightforward to report a relative likelihood of one response over another. In these cases, we report only the direction of the change, not a relative likelihood.

In all cases, we indicate the independent variables included in the model.

Type I and Type II errors

Due to the sheer number of analyses being run, the likelihood of making a Type I error is increased.[4] We mitigate the number of Type I errors using the Bonferroni correction to compute a lower threshold for statistical significance.[5] In order to be as conservative as possible, we set α lower.[6] We identified the largest number of questions that could be classified as a family (22) and used this to calculate a family-wise error correction: 0.05/22 = 0.002. Using a lower α increases the probability of making a Type II error.[7] Thus, we consider there to be sufficient evidence to suggest a statistically significant difference if the p-value for the test statistic is < 0.002.

We focus on multivariate analyses because they do account for confounding factors. Thus, differences in a multivariate model cannot be explained by factors in the model. That

[4] A Type I error is incorrectly rejecting a true null hypothesis; for more information on Type I errors in statistical hypothesis testing, see [1].

[5] For more information on the Bonferroni correction, see [5].

[6] α is set for a hypothesis test as the threshold for making a Type I error.

[7] A Type II error is failing to reject a false null hypothesis, thinking there is no difference when there really is; for more information on Type II errors in statistical hypothesis testing, see [1].

is, if we see a statistically significant difference between the responses of men and women in a multivariate model that includes discipline, geographic region, country development level, and employment sector, then we believe that the difference in the men's and women's responses is due to gender and not to any other factors. We ran a series of binary logistic and ordinal logistic regressions with the variable of interest as the dependent variable and the following included as independent (explanatory) variables: gender, age (as a proxy for career progress), discipline, employment sector, geographic region, and level of development.

The types of logistic regressions differ because many of our dependent variables are binary (for example, Yes/No). Some are ordinal, (for example, Strongly agree / Agree / Neutral / Disagree / Strongly Disagree), and for these, we combined Strongly agree and Agree into "agree"; we did the same thing with "disagree." Thus, we have: agree / neutral / disagree.

The regression coefficients can be interpreted to tell us whether one group is more (or less) likely than another group to respond in a particular way while accounting for potential confounding factors. So, we can test for differences between women's and men's responses to a question while accounting for potential confounding factors (age, discipline, employment sector, geographic region, and level of development). Similarly, we can test for differences across disciplines while accounting for potential confounding factors (gender, age, employment sector, geographic region, and level of development).

2 Tools Produced

The questionnaire is available online at:
https://statisticalresearchcenter.aip.org/global18.

3 Results

There were 32,346 respondents to the first question:

> *"Have you studied or worked in mathematical, computing or natural sciences, or in the history and philosophy of science and technology?"*

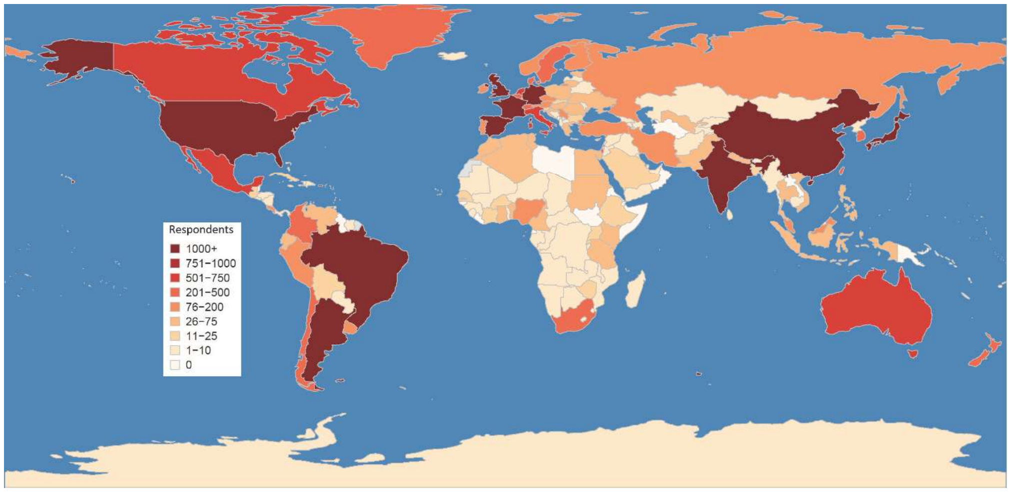


Figure 2: Number of respondents by country.

Table 1: Proportion of women and men among respondents by discipline.

Discipline	Women (%)	Men (%)	n
Astronomy	48	52	2597
Biology	69	31	2960
Chemistry	51	49	2698
Computer Science	55	45	3150
History and Philosophy of Science	46	54	324
Mathematics	43	57	3458
Mathematics – Applied	54	46	2146
Physics	37	63	7570
Overall	**50**	**50**	—

The overall total does not match the 32,346 respondents to question 1 because there are respondents in other disciplines and not every respondent answered the question about their gender. In addition, a respondent could have a university degree in one discipline and be employed in a different discipline. Thus, these numbers are different at different stages in the questionnaire. The data above reflect the discipline corresponding to "primary field of study."

The number of responses to subsequent questions varied based on skip patterns and respondent participation. We collected data from 159 unique countries, although the number of respondents from each varied (Figure 2, p. 46).

Women and men are almost equally represented in the overall data set. Three hundred and eighty respondents preferred not to respond to the question querying their gender, and, therefore, their responses are not included in these analyses, which focus on differences in experiences between women and men. In Table 1 (p. 46), we provide a breakdown of the proportion of men and women among respondents by discipline. In Table 2 (p. 47), we provide a breakdown of the proportion of men and women among respondents by geographic region.

Table 2: Proportion of women and men among respondents by geographic region.

Geographic Region	Woman (%)	Men (%)	*n*
Africa	61	39	1265
Northern America	56	44	5003
Caribbean and Central America	52	48	664
South America	53	47	3314
Western Asia	59	41	395
Central and Southern Asia	45	55	1416
Eastern and South-eastern Asia	36	64	4655
Northern Europe	58	42	2097
Western Europe	46	54	7639
Eastern Europe	47	53	574
Southern Europe	53	47	2039
Oceania	60	40	927
Overall	**50**	**50**	—

The overall total does not match the 32,346 respondents to question 1 because not every respondent answered the question about their gender. In addition, a respondent could have graduated from secondary school in one geographic region, earned their first university degree in a second, earned a second university degree in another, and be primarily employed in yet another. Thus, these numbers vary at different stages in the questionnaire. The data above are for the respondents' current location.

3.1 Multivariate Analysis: Sexual Harassment

We first turn to the issue of harassment. We asked a question: Have you ever encountered sexual harassment at school or work? We compare the respondents who replied *"yes, it happened to me"* with those who did not reply that *"it happened to me."* We first ran a model with all respondents with gender, age, discipline, employment sector, geographic region, and Human Development Index as independent variables.

Results from this model allow us to compare the experiences of men and women while accounting for age, discipline, employment sector, geographic region, and level of human development. The results indicate that women are 14.4 times more likely to indicate having personally experienced sexual harassment at school or work than men – 1440% more likely.

We then examined the experience of women in each discipline separately using gender, age, employment sector, geographic region, and HDI as independent variables.

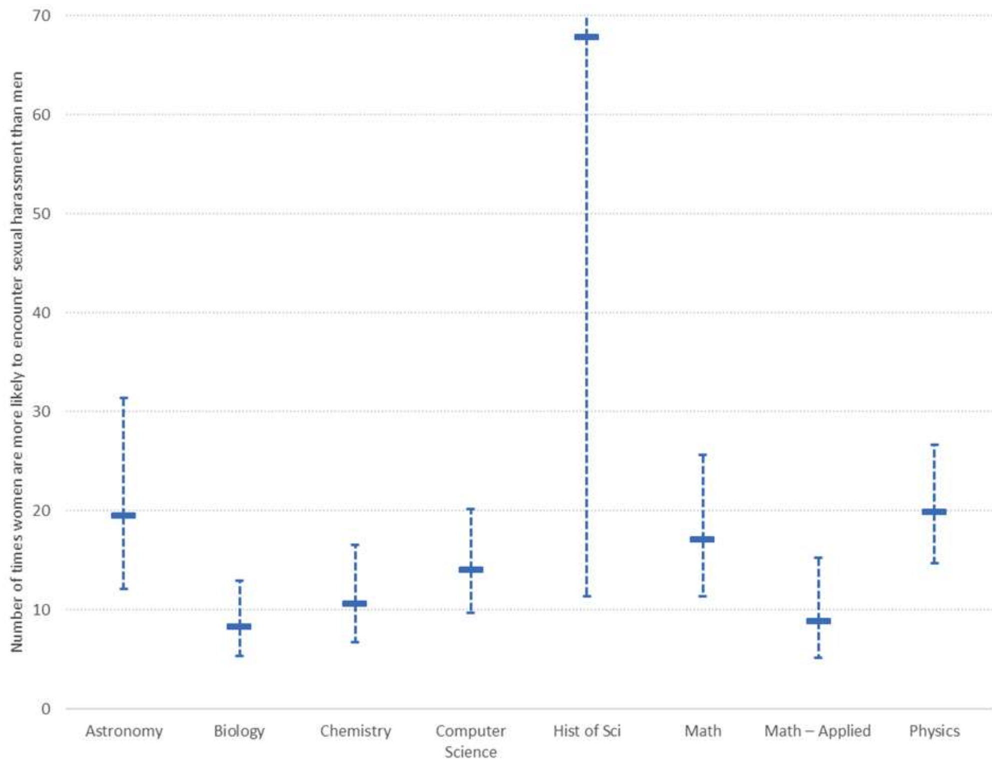
The results are shown graphically in Figure 3 (p. 49). In short, women in every discipline were more likely to indicate having personally experienced sexual harassment at school or work than men. Given the standard error in the estimates, we can say there is no one discipline in which the likelihood is lower for women than any other discipline.

While these results are consistent with the observations in the bivariate analyses (Figure 4, p. 50), the raw data from Figure 4 (p. 50) *understates* the relative likelihood for women, compared to men, to have experienced sexual harassment at school or work after taking into account confounding factors such as age, geographic region, employment sector, and level of development.

The data in Figure 3 (p. 49) should be used to measure or quantify the relative likelihood for women compared to men. For example, according to Figure 3 (p. 49), a woman in astronomy is about 20 times more likely to encounter sexual harassment at school or work than a man. Yet, the raw data (Figure 4, p. 50) shown in the bivariate (30% for women versus 3% for men) *understates* relative likelihood in astronomy. Furthermore, the relative value in Figure 4, p. 50 (30%/3% = 10 times more likely) is below the lower point of the interval estimated based on the multivariate model.

We also examined the experience of women in each geographic region separately using gender, age, discipline, employment sector, and HDI as independent variables.

The results are shown graphically in Figure 5 (p. 51). In short, women in every geographic region were more likely to indicate having personally experienced sexual harassment at



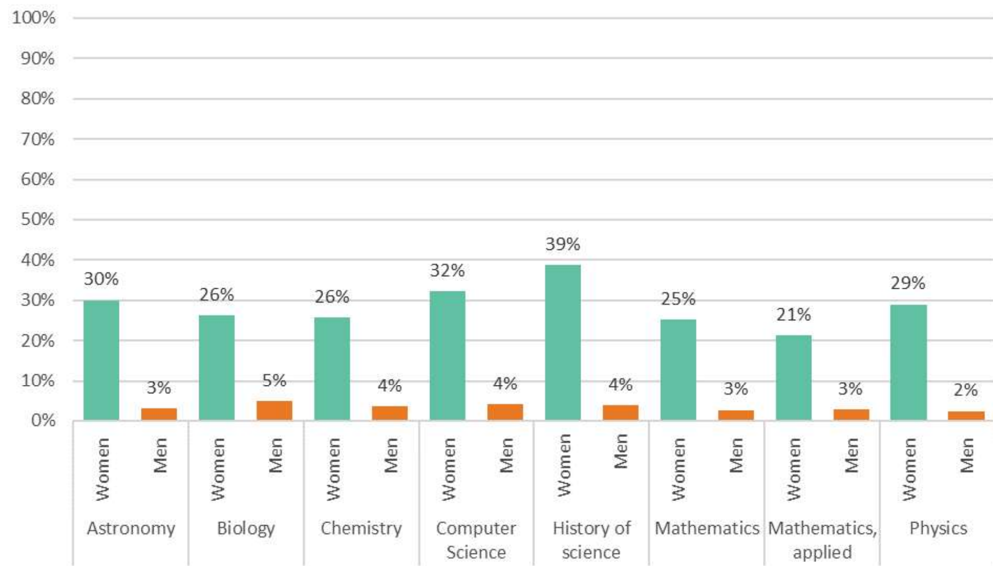
The horizontal bar indicates the point estimate, and the dashed lines indicate the standard error. The standard error is much larger for History of Science due to a smaller number of respondents.

Figure 3: Encountering sexual harassment at school or work by discipline.

school or work than men, except for Western Asia. This means that there was no statistical difference between men and women in Western Asia. Given the standard error in the estimates, we can say there is no one geographic region in which the likelihood is lower for women than any other geographic region. Again, this is consistent with the bivariate findings (Figure 6, p. 52), but the bivariate image *understates* the relative likelihood in most cases.

We also examined the experience of women by *level of development*[8] separately using gender, age, discipline, employment sector, and geographic region as independent variables.

[8]Higher development includes countries with an HDI 0.700 or greater; lower development includes those with and HDI < 0.700.



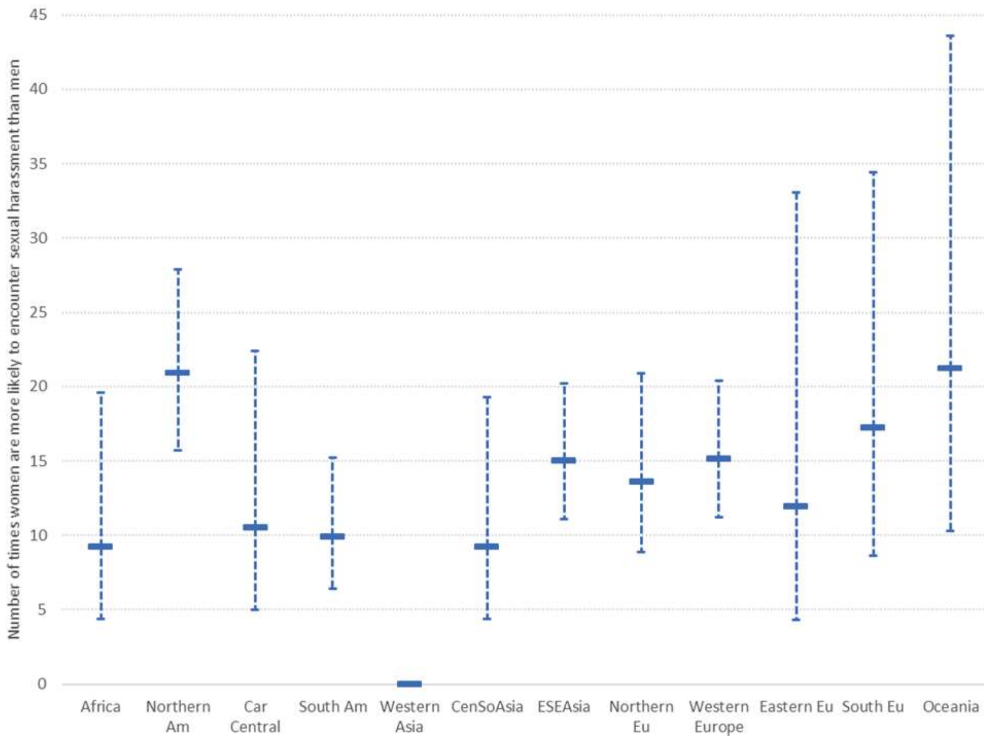
Each of these statistically significant relative differences (30% to 3%, or 10 times, in Astronomy, for example) *understates* the relative likelihood between women and men in the multivariate model which accounts for confounding factors including age, geographic region, employment sector, and level of development.

Figure 4: Respondents indicating they personally encountered sexual harassment at school or work by discipline.

The results are shown graphically in Figure 7 (p. 53). In short, women in both HDI groups were more likely to indicate having personally experienced sexual harassment at school or work than men, Given the standard error in the estimates, we can say there is no difference in the likelihood for women by HDI. Again, this is consistent with the findings in the bivariate analysis (see Figure 8, p. 54).

Finally, we examined the experience of women in different employment sectors separately using gender, age, discipline, geographic region, and HDI as independent variables.

The results are shown graphically in Figure 9 (p. 55). In short, women in every employment sector were more likely to indicate having personally experienced sexual harassment at school or work than men. Given the standard error in the estimates, we can say there is no employment sector in which the likelihood is lower for women than any other employment sector.

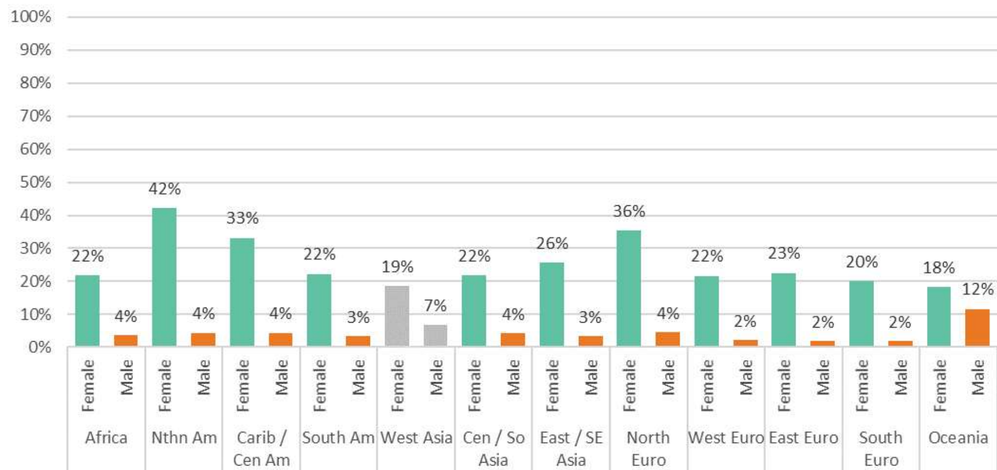


The horizontal bar indicates the point estimate, and the dashed lines indicate the standard error. The size of the standard error depends on the number of respondents.

Figure 5: Encountering sexual harassment at school or work by geographic region.

The statistical models are not the only source of information. We have begun conducting qualitative analysis on some open-ended questions. One respondent reported:

> *"My institution does not have a defined sexual harassment policy. I reported a supervisor for sexual harassment and had a terrible experience during the HR investigation."*



Each of these statistically significant relative differences (22% to 4%, or 5.5 times, in Africa, for example) *understates* the relative likelihood between women and men in the multivariate model which accounts for confounding factors including age, discipline, employment sector, and level of development.

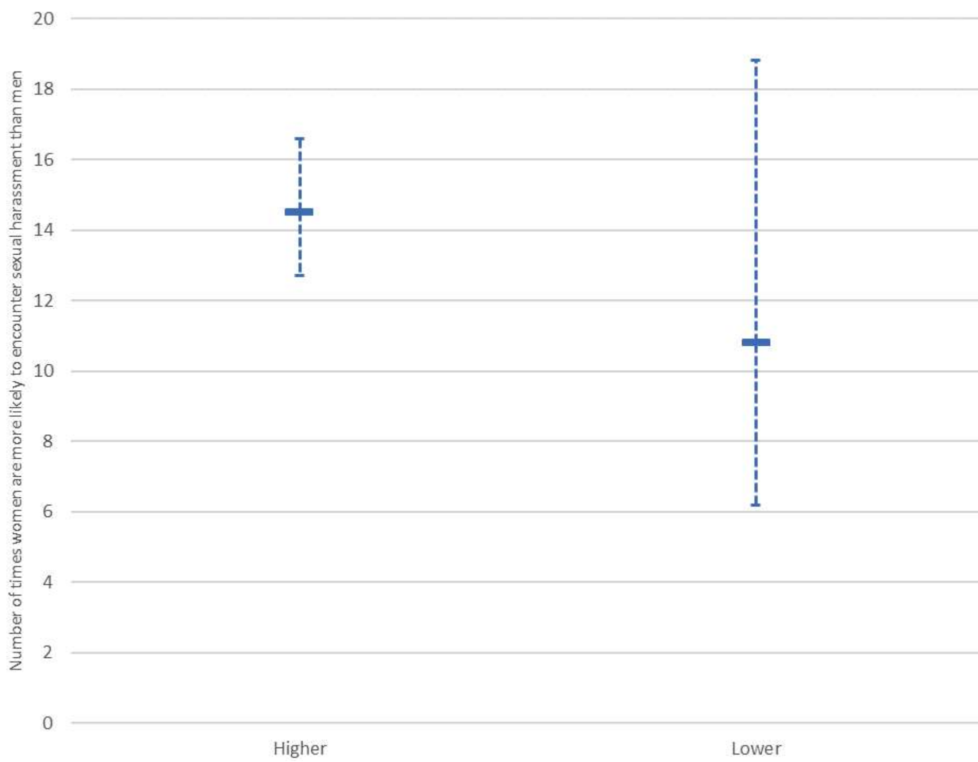
The gray bars indicate that the difference is not statistically significant at the 0.002 level.

Figure 6: Percentages of respondents indicating they personally encountered sexual harassment at school or work by geographic region.

3.2 Multivariate Analysis: Gender Differences in Educational Experiences

We now consider respondents' educational experiences. We turn now to the multivariate analysis, which can account for confounding factors that the bivariate analysis cannot reveal. We ran logistic regressions for sixteen questions from the survey:

- *Choose field.* When did you choose your primary field of study?
- Who encouraged you in your studies?
  - *Encourage: Partner.* Spouse or partner
  - *Encourage: Parents.* Parents
  - *Encourage: Other family.* Other family members
  - *Encourage: Employ/colleagues.* Employer or colleagues
  - *Encourage: Community.* Your neighborhood, community, or friends
  - *Encourage: Teachers/Profs/Mentors.* Teachers, professors, or mentors

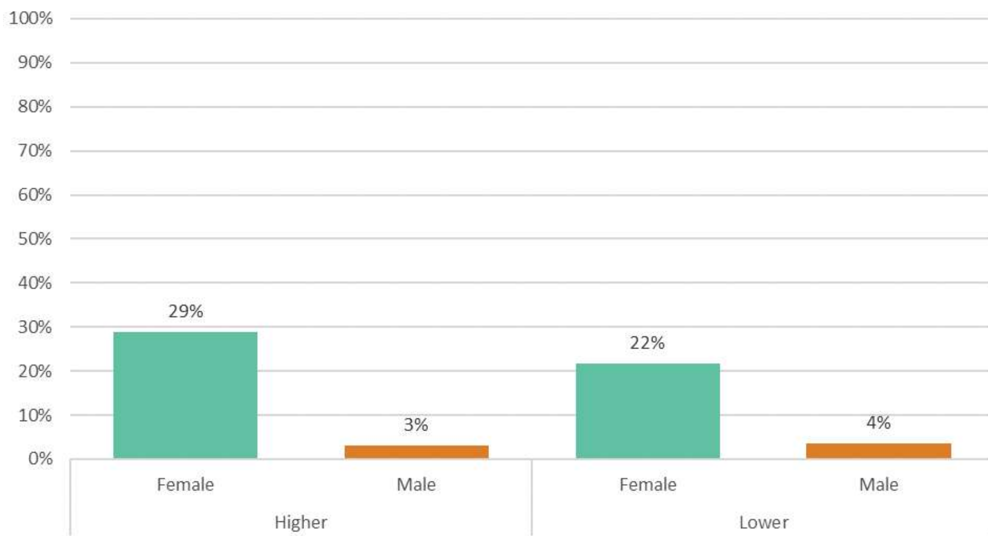


The horizontal bar indicates the point estimate, and the dashed lines indicate the standard error. The size of the standard error depends on the number of respondents.

Figure 7: Encountering sexual harassment at school or work by level of development.

- *Encourage: Other students.* Other students
- *Encourage: Own determination.* Your own determination, will power, and hard work

- *Doctoral program quality.* How would you rate the quality of your doctoral program?
- *Doctoral advisor support.* In my doctoral experience, I had support from my advisor or supervisor.
- *Doctoral program fair treatment.* In my doctoral experience, my program treated everyone fairly.
- *Interruptions in doctoral studies.* Have there been any significant interruptions in your doctoral studies?

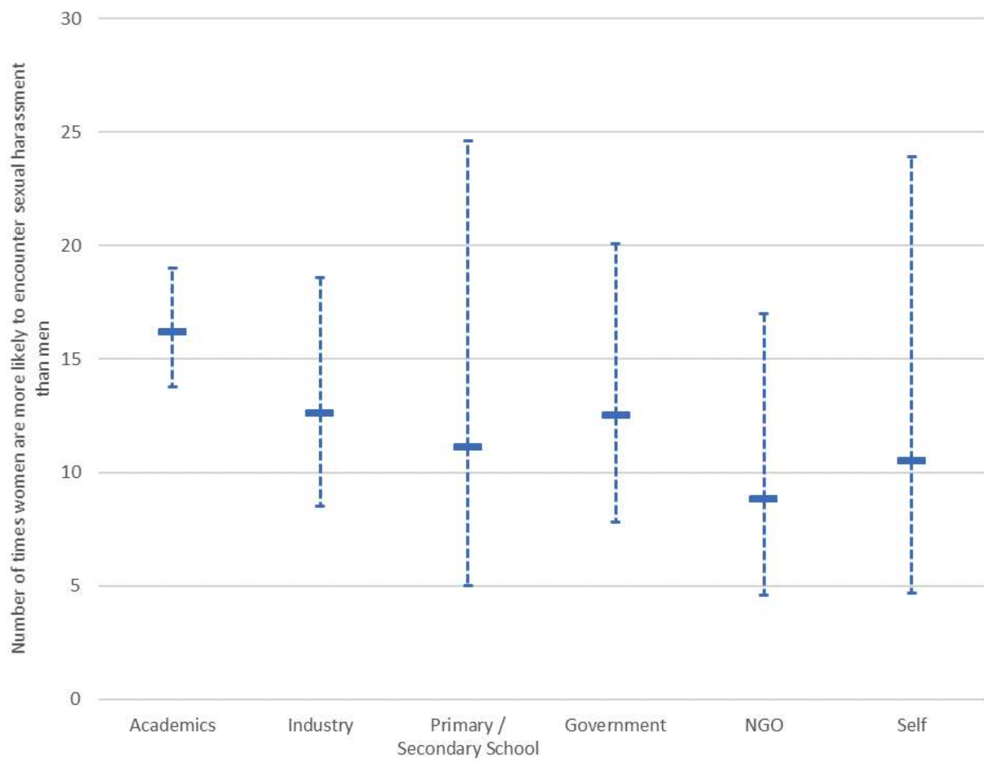


Both of these statistically significant relative differences 29% to 3%, or ~ 10 times, in countries with higher development levels, for example) *understate* the relative likelihood between women and men in the multivariate model which accounts for confounding factors including age, discipline, employment sector, and geographic region.

Figure 8: Percentages of respondents indicating they personally encountered sexual harassment at school or work by level of human development.

We ran multivariate models for each of these twelve questions using gender, age, discipline, geographic region, and HDI as independent variables. Note that we do not include employment sector because the variables of interest are about educational experiences. The results are shown in Table 3 (p. 56).

Overall, we find no statistically significant gender differences in the timing of choosing one's discipline. Women are more likely to report that their families encouraged them in their studies than men, and women are more likely than men to report being encouraged by their own determination, will power, and hard work. There are no statistically significant differences in the responses of men and women regarding other sources of encouragement. Men are more likely to report positive experiences in their doctoral programs, and women are more likely to report interruptions in their doctoral programs. One woman who responded told us,



The horizontal bar indicates the point estimate, and the dashed lines indicate the standard error. The size of the standard error depends on the number of respondents.

Figure 9: Encountering sexual harassment at school or work by employment sector.

> *"Unconditional support from my family and the available resources form my parents"*

had helped her succeed. Another attributed her success to

> *"being stubborn and unwilling to let the bad guys win!!!"*

On the other hand, another respondent told us about her thesis advisor:

> *"Harassment (psychological) by my thesis director, which was persistent, and people knew about it in the laboratory, but [I did not receive] any support for fear of reprisals."*

These echo the findings above.

A Closer Look at Gender Differences in Educational Experiences by Discipline

In order to better understand gender differences by discipline, we ran a multivariate model for each discipline separately. That allows us to look at gender differences within a specific discipline. We examine some of those findings below. We find that there is a statistically significant difference in the timing for men and women whose primary field is computer science or mathematics – applied (see Figure 10, p. 57).

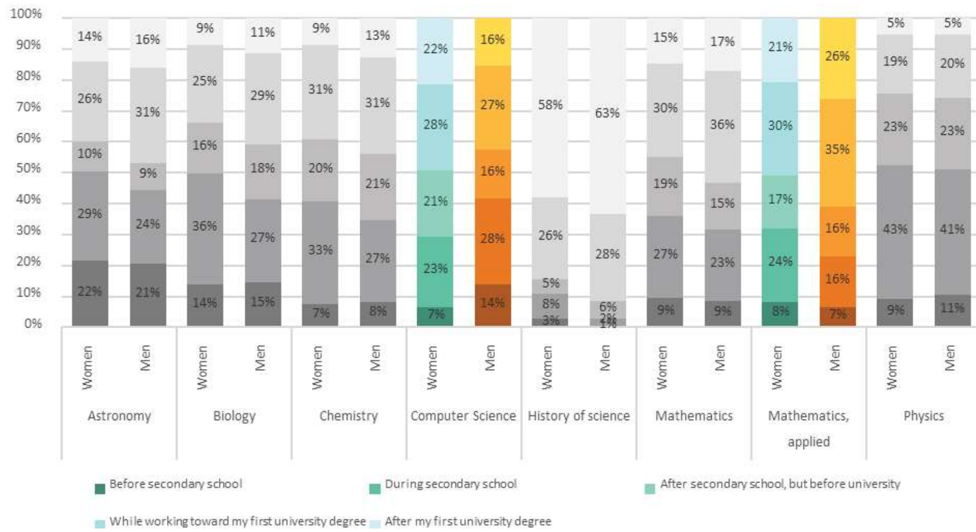
We also find that there is a statistically significant difference in the rating of doctoral program quality for men and women whose primary field is astronomy, chemistry, math-

Table 3: Gender differences in educational experiences.

Item	**Gender Differences**
Choose field	No statistically significant gender differences.
Encourage: Partner	**Women 1.7 times more likely** to say **spouse or partner** than men.
Encourage: Parents	**Women 1.4 times more likely** to say **parents** than men.
Encourage: Other family	**Women 1.2 times more likely** to say **other family members** than men.
Encourage: Employ / colleagues	No statistically significant gender differences.
Encourage: Community	No statistically significant gender differences.
Encourage: Teachers / profs / mentors	No statistically significant gender differences.
Encourage: Other students	No statistically significant gender differences.
Encourage: Own determination	**Women 1.1 times more likely** to say **own determination** than men.
Doctoral program quality†	**Men likely** to rate quality **higher** than women.
Doctoral advisor support†	**Men likely** to rate advisor support **better** than women.
Doctoral program fair treatment†	**Men likely** to **agree more strongly** than women.
Interruptions in doctoral studies	**Women 1.6 times more likely** to say **Yes** than men.

These gender differences are statistically significant at the $\alpha = 0.002$ level after accounting for potential confounding factors (age, discipline, geographic region, and level of development).

†The dependent variable is ordinal, so we are unable to report a single odds ratio.



Gray indicates the differences are not statistically significant by gender at the 0.002 level.

Figure 10: Percentages of respondents indicating the period of time they chose their primary field by discipline.

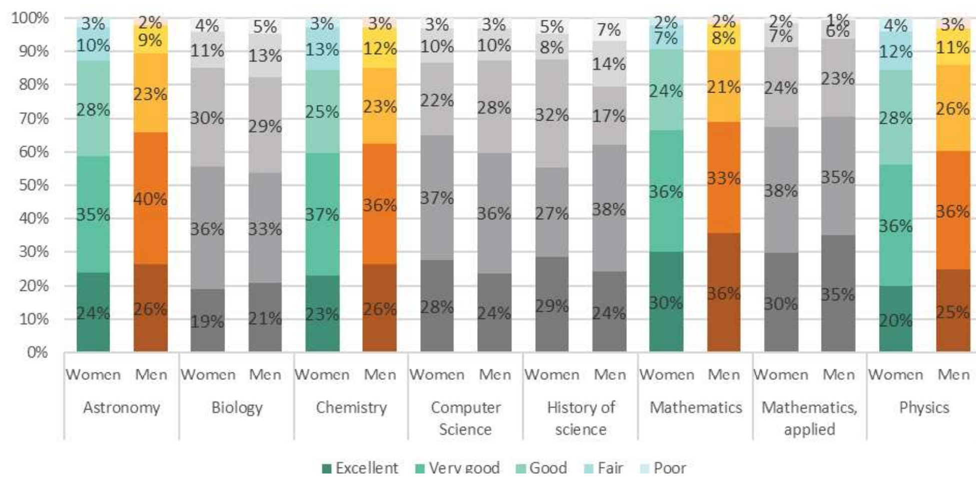
ematics, or physics (see Figure 11, p. 58). We see a similar pattern for the respondents indicating they had a positive relationship with their advisor or supervisor during doctoral studies (see Figure 12, p. 58).

We also see statistically significant differences in respondents' opinion of fair treatment in their doctoral program by gender (Figure 13, p. 59) and in respondents indicating they had a significant interruption in their doctoral studies by discipline (Figure 14, p. 59).

A Closer Look at Gender Differences in Educational Experiences by Geographic Region

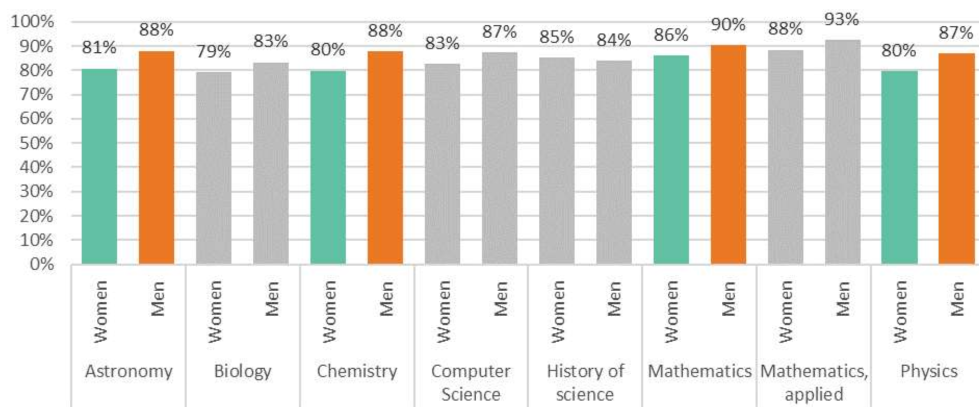
As we did with each of the disciplines, we ran a multivariate model for each geographic region separately. We examine some of those findings below. We find that there is a statistically significant difference in the proportion of respondents indicating they had a positive relationship with their advisor or supervisor during doctoral studies (see Figure 15, p. 60).

We also find that there is a statistically significant difference in the rating of fair treatment in the doctoral program for men and women who studied in several regions (see Figure 16,



Gray indicates the differences are not statistically significant by gender at the 0.002 level.

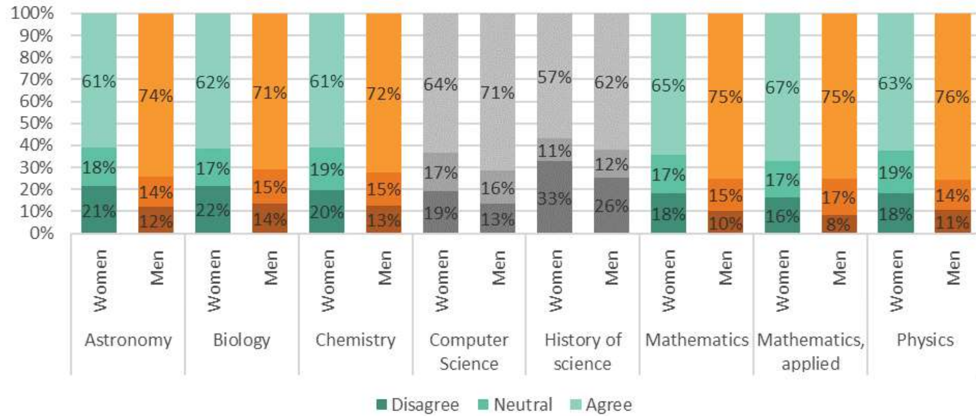
Figure 11: Percentages of respondents' rating of quality of doctoral program by discipline.



Gray indicates the differences are not statistically significant by gender at the 0.002 level.

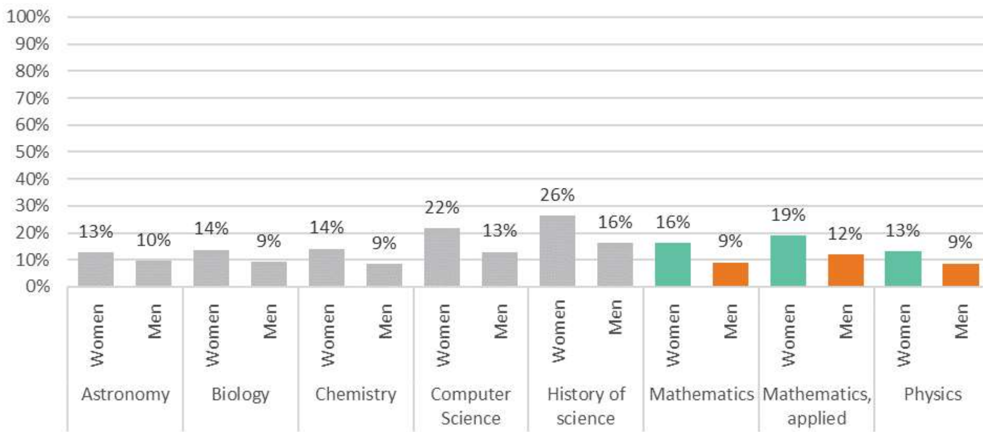
Figure 12: Percentages of respondents indicating they had a positive relationship with their advisor or supervisor during doctoral studies by discipline.

p. 60). Figure 17 (p. 61) depicts gender differences in interruptions in doctoral programs by geographic region.



Gray indicates the differences are not statistically significant by gender at the 0.002 level.

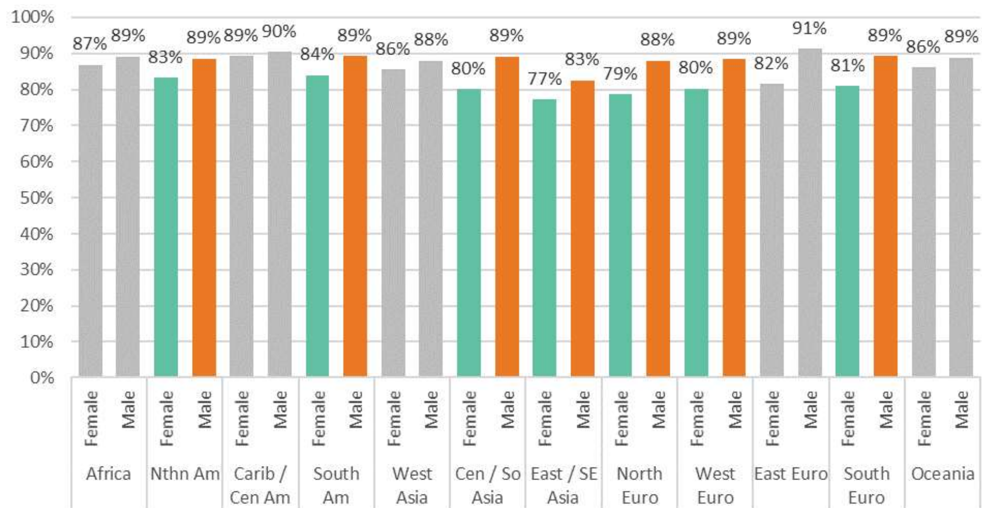
Figure 13: Percentages of sespondents' agreement with the statement, *"My program treated everyone fairly"* by discipline.



Each of these statistically significant relative differences (16% for women to 9% for women, or ~ 1.8 times, in Mathematics, for example) *understates* the relative likelihood between women and men in the multivariate model which accounts for confounding factors including age, employment sector, geographic region, and level of development.

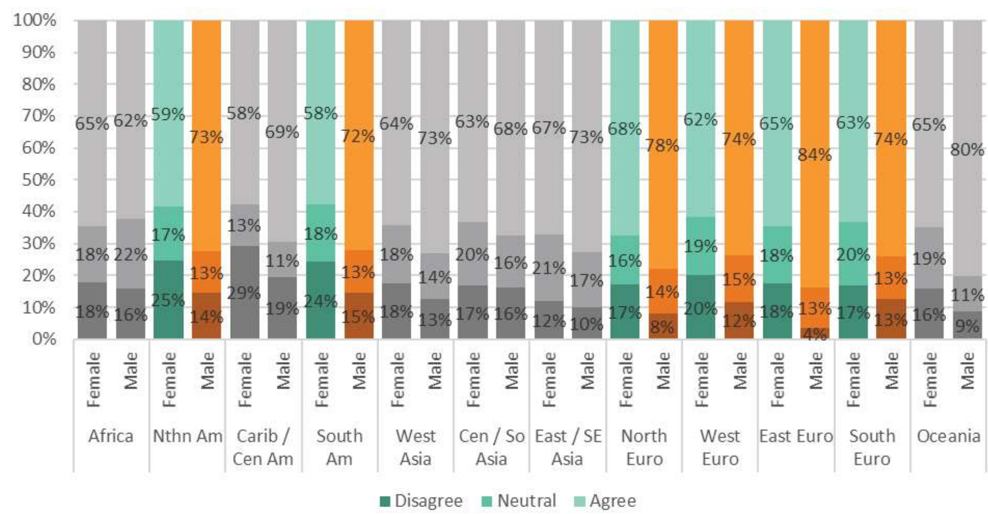
Gray indicates the differences are not statistically significant by gender at the 0.002 level.

Figure 14: Percentages of respondents indicating they had dignificant interruptions in their doctoral studies by discipline.



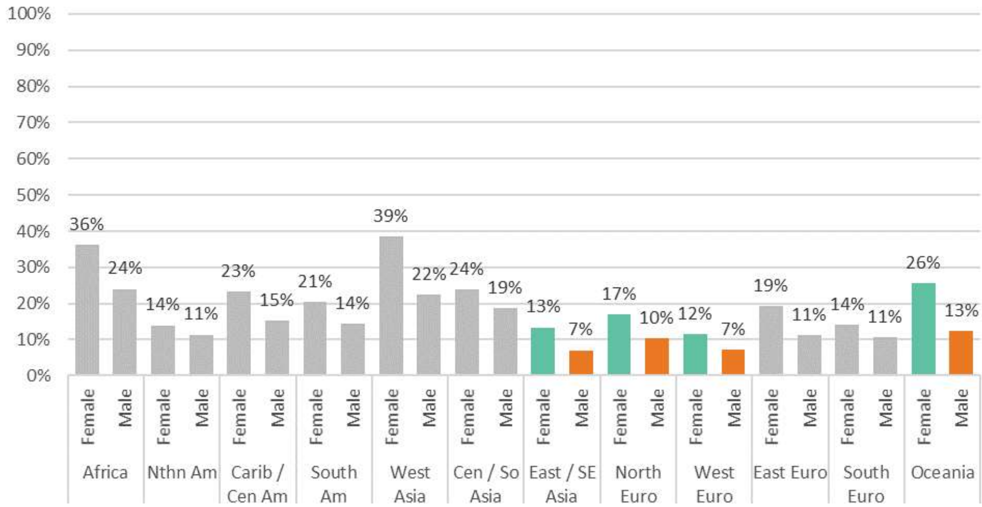
Gray indicates the differences are not statistically significant by gender at the 0.002 level.

Figure 15: Percentages of respondents indicating they had a positive relationship with their advisor or supervisor during doctoral studies by geographic region.



Gray indicates the differences are not statistically significant by gender at the 0.002 level.

Figure 16: Percentages of respondents agreement with the statement, *"My program treated everyone fairly"* by geographic region.



Each of thesestatistically significant relative differences (13% for women to 7% for men, or ~ 1.9 times, in East / SE Asia, for example) *understates* the relative likelihood between women and men in the multivariate model which accounts for confounding factors including age, employment sector, discipline, and level of development.

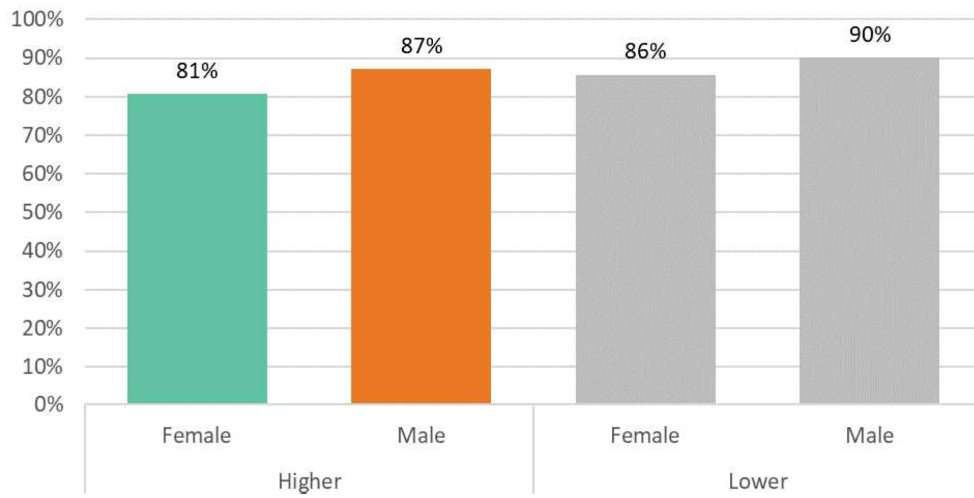
Gray indicates the differences are not statistically significant by gender at the 0.002 level.

Figure 17: Percentages of respondents indicating they had significant interruptions in their doctoral studies by geographic region.

A Closer Look at Gender Differences in Educational Experiences by Level of Development

As we did with discipline and geographic region, we ran separate multivariate models for countries with higher and lower levels of economic development. We examine some of those findings below. We find that there is no statistically significant difference in the ratings of doctoral program quality by level of development. We do see a statistically significant difference in the proportion of respondents indicating they had a positive relationship with their advisor or supervisor during doctoral studies in countries with higher levels of development, but not in those with a lower level of development (see Figure 18, p. 62).

We also find that there is a statistically significant difference in the rating of fair treatment in the doctoral program for men and women whose primary field in several countries with



Gray indicates the differences are not statistically significant by gender at the 0.002 level.

Figure 18: Percentages of respondents indicating they had a positive relationship with their advisor or supervisor during doctoral studies by level of development.

higher levels of economic development (see Figure 19, p. 63). Figure 20 (p. 64) depicts gender differences in interruptions in doctoral programs by level of development.

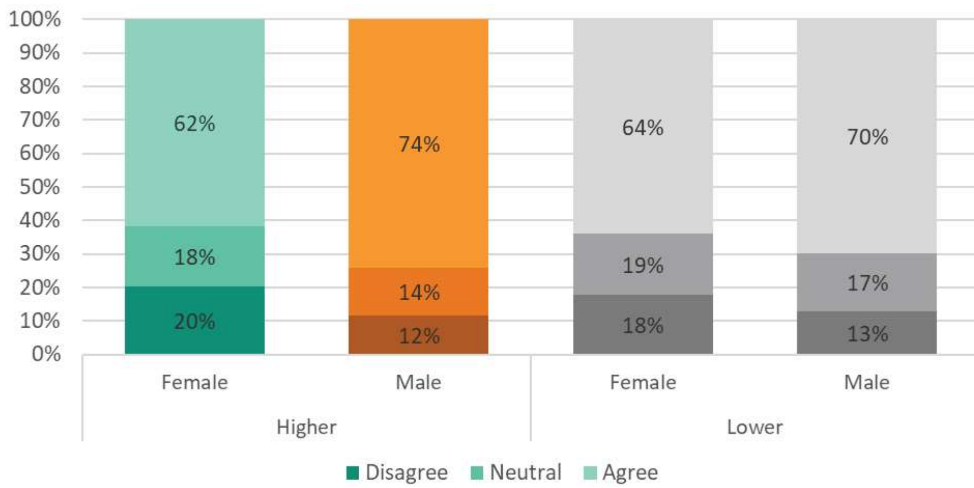
3.3 Multivariate Analysis: Gender Differences in Employment Experiences

We have seen that there are gender differences in educational experiences. We next tested for any gender differences in employment experiences. As before, we ran a series of multivariate regression models examining responses to the following questions:

- *Employer fair treatment.* At my current job, my employer treats everyone fairly.
- *Respectful co-workers.* My co-workers are respectful of everyone.
- *Rate of career progression.* Compared to colleagues who completed their final degrees at the same time as you, how quickly have you progressed in your career?

Our multivariate models used gender, age, discipline, employment sector, geographic region, and HDI as independent variables. The results are shown in Table 4 (p. 63).

In each of these items, men reported more positive experiences than women. One noted:



Gray indicates the differences are not statistically significant by gender at the 0.002 level.

Figure 19: Percentages of respondents' agreement with the statement, *"My program treated everyone fairly"* by level of development.

> *"Sexism is constant and prevalent everywhere, and it's exhausting."*

In some situations, women feel isolated in addition to facing the more negative experiences. One respondent wrote:

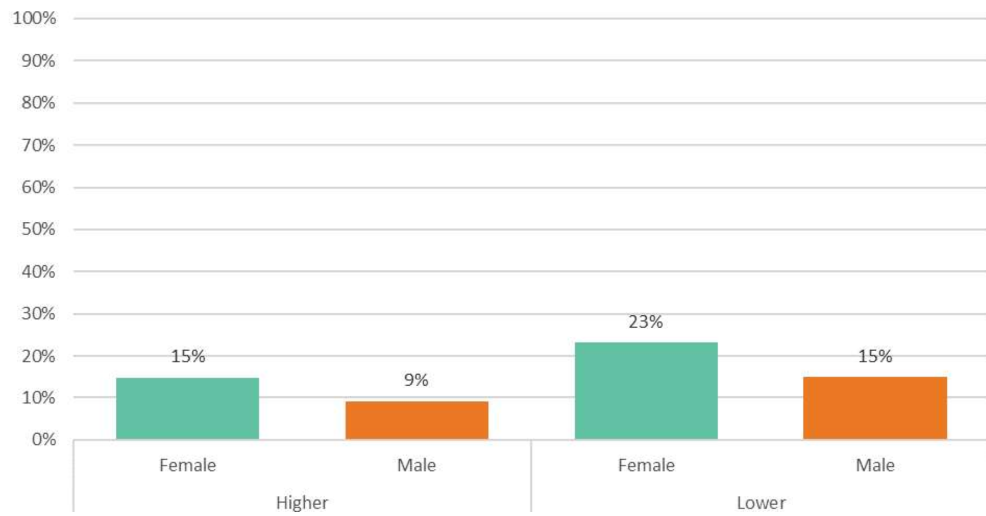
> *"Working in a fully male environment, I sometimes feel there is no one to talk to."*

Table 4: Gender differences in employment experiences.

Item	Gender Differences
Employer fair treatment†	**Men likely** to **agree more strongly** than women.
Respectful co-workers†	**Men likely** to **agree more strongly** than women.
Rate of career progression†	**Men likely** to report **faster career progression** than women.

These gender differences are statistically significant at the α = 0.002 level after accounting for potential confounding factors (age, discipline, geographic region, and level of development).

†The dependent variable is ordinal, so we are unable to report a single odds ratio.



Both of these relative differences (15% for men to 9% for women, or ~ 1.7 times, in countries with higher development levels, for example) *understate* the relative likelihood between women and men in the multivariate model which accounts for confounding factors including age, employment sector, discipline, and geographic region.

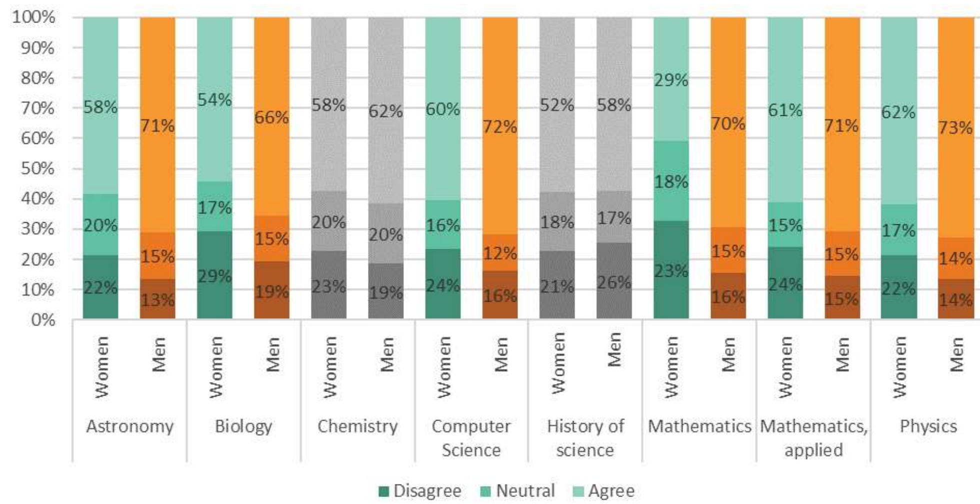
Figure 20: Percentages of respondents indicating they had significant interruptions in their doctoral studies by level of development.

Even when other women are present, women can still feel excluded. Another respondent said:

> *"There are few women in my department, and for the first three years I almost always ate lunch alone. Only the men arrange to eat lunch as a group, and they rarely invite me."*

A Closer Look at Gender Differences in Employment Experiences by Discipline

As before, we ran a multivariate model for each discipline separately to look at gender differences within a specific discipline. We examine some of those findings below. We find that there is a statistically significant difference men's and women's agreement with the statement, *"My employer treats everyone fairly."* (see Figure 21, p. 65).



Gray indicates the differences are not statistically significant by gender at the 0.002 level.

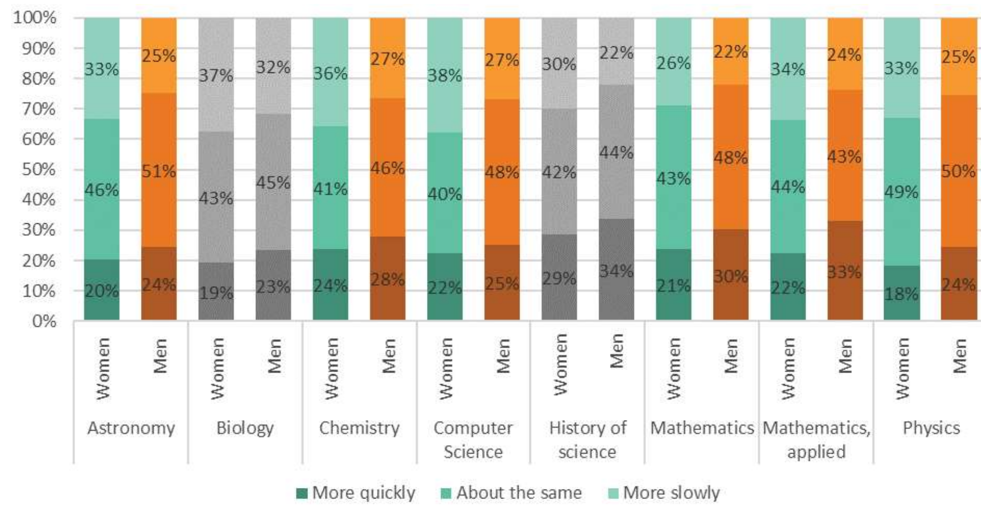
Figure 21: Percentages of respondents' agreement with the statement, *"My employer treats everyone fairly"* by discipline.

We also find that there is a statistically significant difference in respondents' career progression by discipline (see Figure 22, p. 66).

A Closer Look at Gender Differences in Employment Experiences by Geographic Region

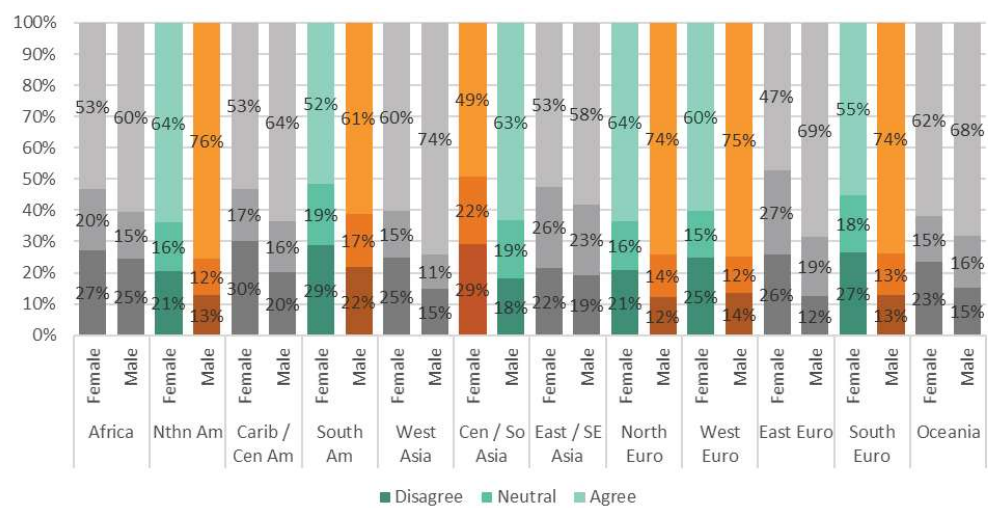
As before, we ran multivariate models for each geographic region in order to better understand gender differences in employment experiences in the different regions. We examine some of those findings below. We find that there is a statistically significant difference in the proportion of respondents indicating their employer treats every fairly (see Figure 23, p. 66).

In regions other than Africa, Caribbean / Central America, and West Asia, men and women had statistically significant difference in their agreement with the statement, *"My co-workers are respectful of everyone."* (see Figure 24, p. 67).



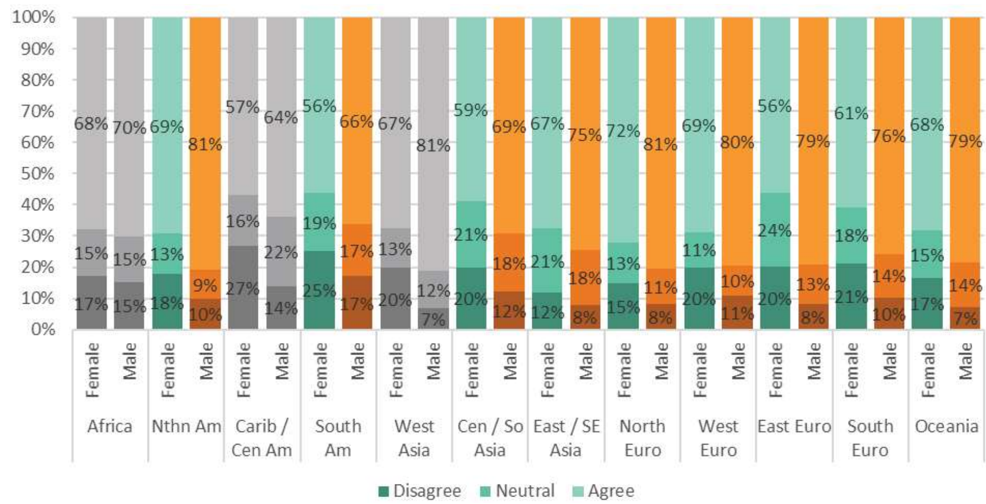
Gray indicates the differences are not statistically significant by gender at the 0.002 level.

Figure 22: Percentages of respondents comparing their career progression to their colleagues who completed their final degrees at the same time by discipline.



Gray indicates the differences are not statistically significant by gender at the 0.002 level.

Figure 23: Percentages of respondents agreement with the statement, *"My employer treats everyone fairly"* by geographic region.



Gray indicates the differences are not statistically significant by gender at the 0.002 level.

Figure 24: Percentages of respondents agreement with the statement, *"My co-workers are respectful of everyone"* by geographic region.

A Closer Look at Gender Differences in Employment Experiences by Level of Development

Again, we ran two multivariate models for countries that with higher and lower levels of economic development to look at gender differences by level of development. We find respondents' agreement with the statement, *"My employer treats everyone fairly,"* is statistically significant difference in both countries with higher and lower levels of development (see Figure 25, p. 68). We see the same pattern with respondents' agreement with the statement, *"My co-workers are respectful of everyone."* (see Figure 26, p. 68).

3.4 Multivariate Analysis: Gender Differences in Access to Resources and Opportunities

We asked questions about respondents' access to resources needed to conduct science and to opportunities to participate in the scientific enterprise. For resources, we asked respondents whether they had *"enough of the following to conduct or present [their] research."*

- Funding

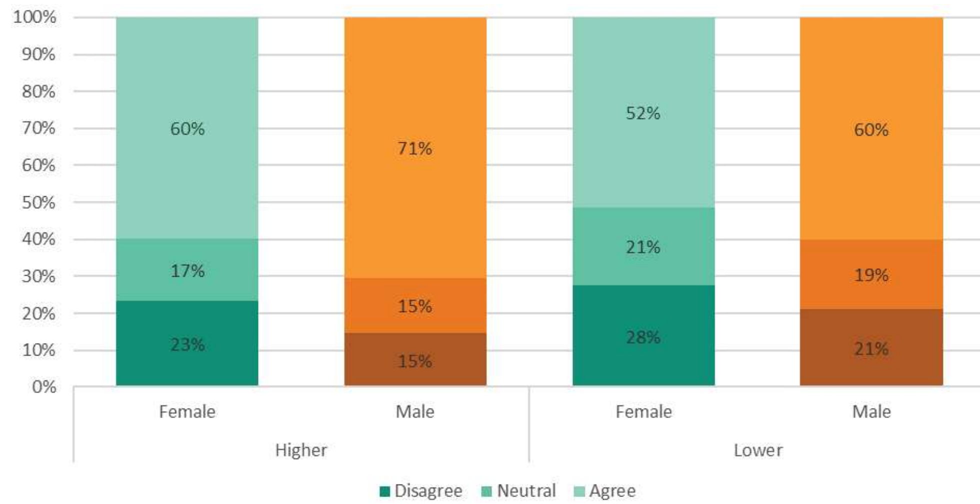


Figure 25: Percentages of respondents' agreement with the statement, *"My employer treats everyone fairly"* by level of development.

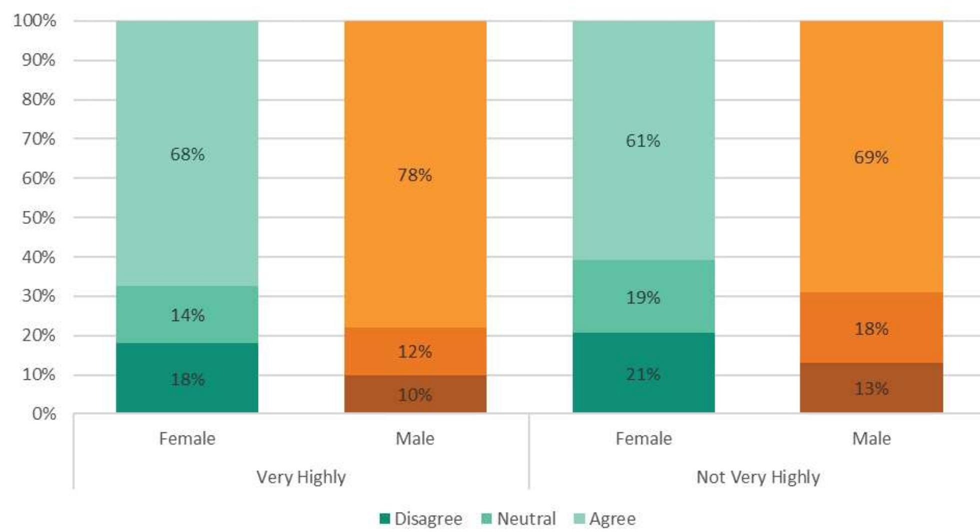


Figure 26: Percentages of respondents' agreement with the statement, *"My co-workers are respectful of everyone"* by level of development.

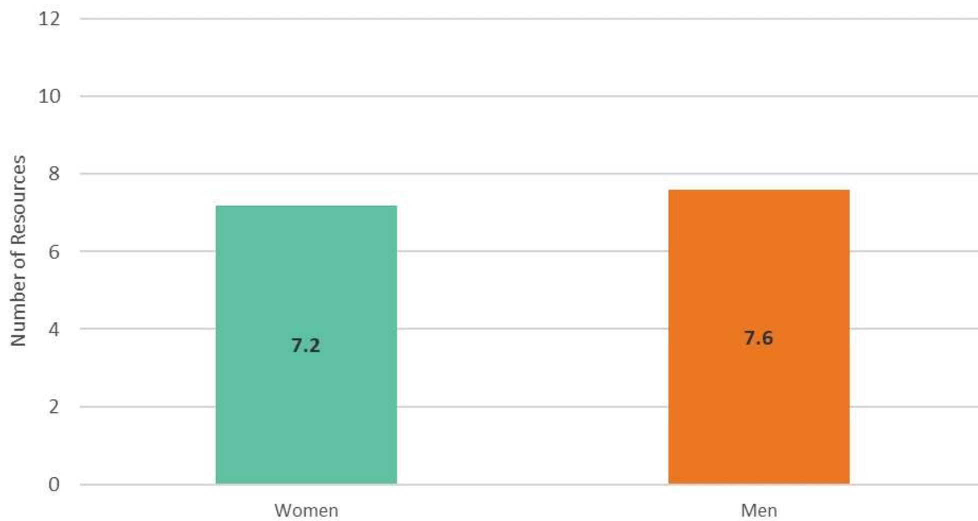
- Office space
- Lab space

- Equipment
- Travel money
- Clerical support
- Employees or students
- Computing capability
- Technical support
- Access to data
- Access to scientific literature
- Support as a working parent

For this analysis of twelve different resources, we summed the number of items for which the respondent indicated *"yes."* Overall, the mean number of resources available to respondents was 7.4 (median = 8). There is a statistically significant difference in the total number of resources available to men (7.6) and women (7.2). However, some of this gender difference might be attributable to differences in discipline, employment sector, geographic region, HDI, or a respondent's age. We ran a multivariate regression to see whether the statistically significant difference in means could be explained by other factors. However, the difference seen in the regression model is virtually the same with men having, on average, 0.4 more resources available than women. While the difference is small, the effect can have a cumulative effect on careers, placing men at a potential advantage compared to women.

For opportunities, we asked respondents whether they had participated in the following:

- Given a talk at a conference as an invited speaker;
- Attended a conference abroad;
- Conducted research abroad;
- Acted as a boss or manager;
- Served as editor of a journal;
- Served on committees for grant agencies;
- Served on important committees at your institute or company;
- Served on an organizing committee for a conference in your field;



This represents access to resources in one year. This difference can compound though one's career.

Figure 27: Average number of resources available to men and women.

- Served on a Board of Directors;
- Was elected to leadership of scholarly associations;
- Advised or supervised undergraduate students;
- Advised or supervised graduate students;
- Served on thesis or dissertation committees (not as advisor or supervisor);
- Given talks or interviews for the general public; for example, TV, newspapers, and magazines.

Out of the fourteen opportunities, respondents reported participating in an average of 6.9. However, the average number of opportunities available to men (7.1) is statistically significantly higher than the average available to women (6.6). As with resources, we recognize that this difference could be due to factors other than gender. We ran a multiple regression model, and we found that discipline, age, employment sector, geographic region, and HDI explain some of the difference. The regression model results indicate that men have, on average, 0.2 more opportunities than women. Again, this is a small difference that can compound over a career.

3.5 Multivariate Analysis: Gender Differences in Discrimination and Career Choices and Progress as a Parent

In the first section, we examined gender differences in sexual harassment, and we found that women are much more likely to experience sexual harassment than men, even after accounting for potential confounding factors. In this section, we look at gender differences in discrimination. We also examine the differences in men's and women's choices about work/life balance and the impact of becoming a parent. Specifically, we examine differences in men's and women's responses to the following questions:

- *Never experienced discrimination.* I have never experienced discrimination.
- *Career influenced relationship decisions.* Has your career influenced your decisions about children, marriage, or a similar long-term partnership?
- *Promotion slowed after children.* My career or rate of promotion slowed significantly when I became a parent.
- *No career change after children.* My work or career did not change significantly when I became a parent.

Men are almost five times more likely never to have experienced discrimination than women. This holds after accounting for potential confounding factors such as age, disci-

Table 5: Gender differences in employment experiences as a parent.

Item	Gender Differences
Never experienced discrimination	**Men 4.8 times more likely** to **never have experienced discrimination** than women.
Career influenced relationship decisions	**Women 1.6 times** more likely to say **Yes** than men.
Promotion slowed after children	**Women 3.3 times more likely** to report **slower** rate of promotion after children than men.
No career change after children	**Men 3.0 times more likely** to report **no change after children** than women.

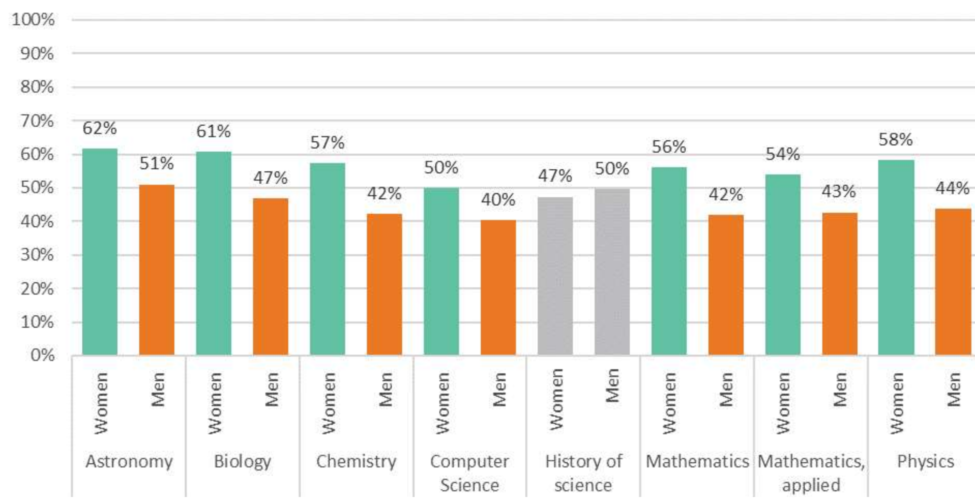
These gender differences are statistically significant at the $\alpha = 0.002$ level after accounting for potential confounding factors (age, discipline, geographic region, and level of development).

pline, employment sector, geographic region, and HDI. Women are also more likely to say that their career had influenced their decisions about marriage and children.

Finally, the effect of children on men's and women's career are vastly different. Women are about three times more likely to report slower rates of promotion after becoming a parent than men. This is consistent with the result that men are about three times more likely to report no change in career progress after becoming a parent.

A Closer Look at Gender Differences in Career Choices and Progress as a Parent by Discipline

As in the previous sections, we ran multivariate models for each discipline. We examine some of those findings below. We find that there is a statistically significant difference between men's and women's responses to whether their career influenced their decisions about children, marriage, or a similar long-term partnership by discipline (see Figure 28, p. 72).

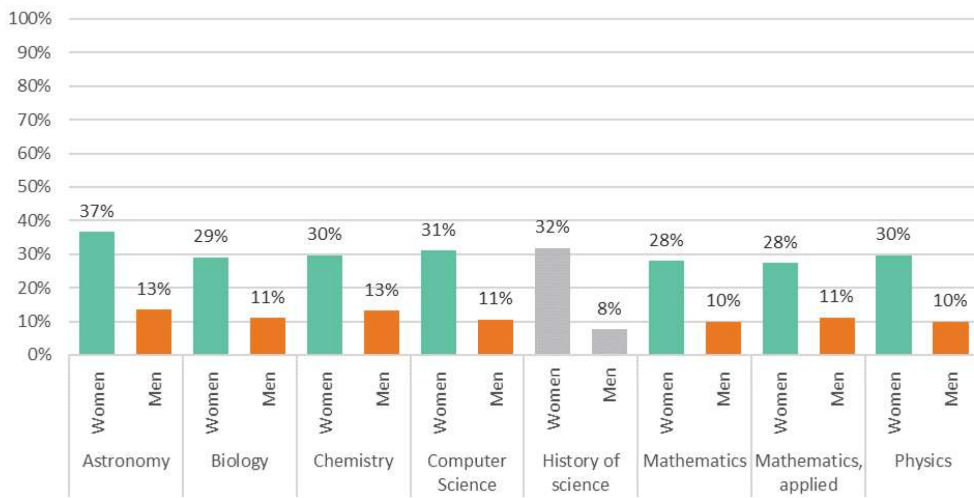


Each of these statistically significant relative differences (62% to 51%, or ~ 1.2 times, in Astronomy, for example) *understates* the relative likelihood between women and men in the multivariate model which accounts for confounding factors including age, employment sector, geographic region, and level of development.

Gray indicates the differences are not statistically significant by gender at the 0.002 level.

Figure 28: Percentages of respondents indicating their career influenced their decisions about children, marriage, or a similar long-term partnership by discipline.

We also find that there is a statistically significant difference in respondents' career or rate of promotion because they became a parent by discipline (see Figure 29, p. 73). A similar pattern emerges for respondents indicating their career had not changed significantly because they became a parent (Figure 30, p. 74).



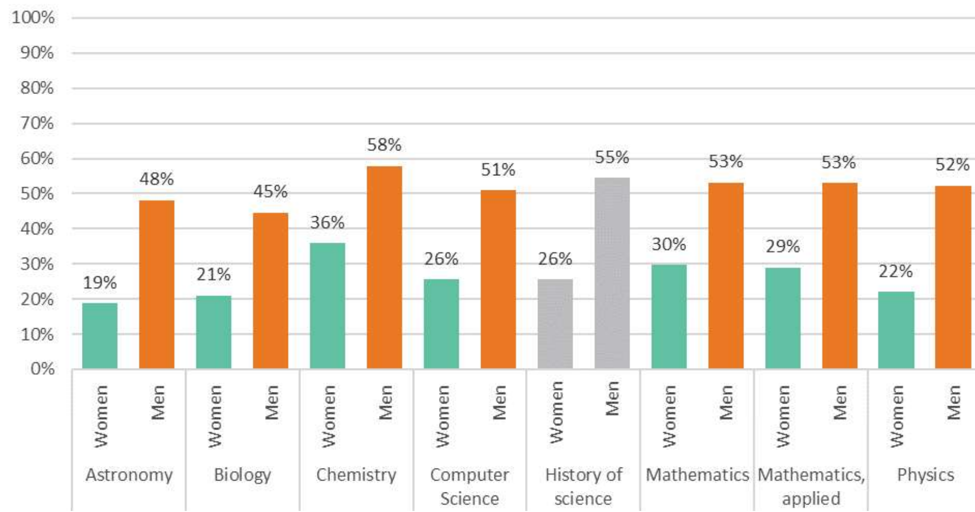
Each of these statistically significant relative differences (37% to 13%, or ~ 2.8 times, in Astronomy, for example) *understates* the relative likelihood between women and men in the multivariate model which accounts for confounding factors including age, employment sector, geographic region, and level of development.

Gray indicates the differences are not statistically significant by gender at the 0.002 level.

Figure 29: Percentages of respondents indicating their Career or rate of promotion slowed significantly because they became a parent by discipline.

A Closer Look at Gender Differences in Discrimination and Career Choices and Progress as a Parent by Geographic Region

Once again, we ran multivariate regression models for each geographic region. Men's careers were less likely to change after becoming a parent in every geographic region (see Figure 31, p. 75).



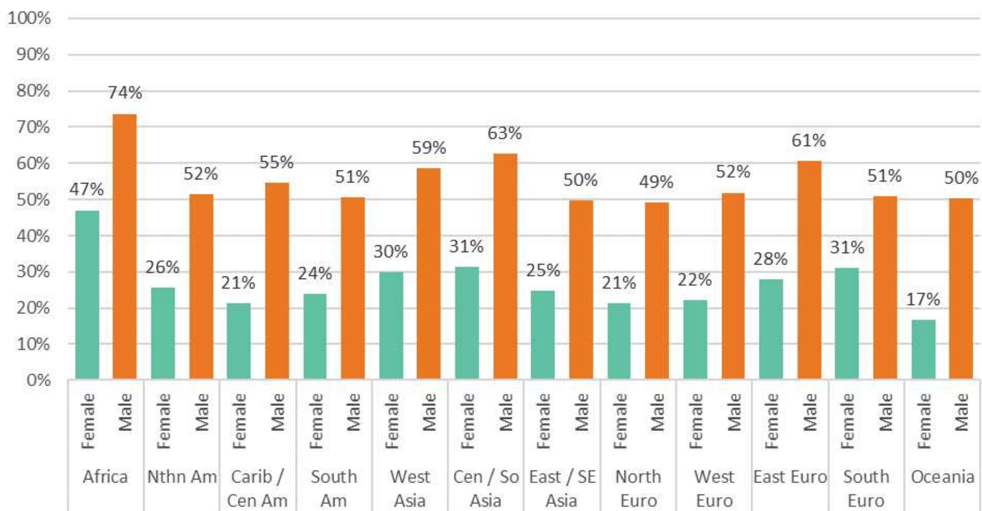
Each of these statistically significant relative differences (48% for men to 19% for women, or ~ 2.5 times, in Astronomy, for example) *understates* the relative likelihood between women and men in the multivariate model which accounts for confounding factors including age, employment sector, geographic region, and level of development.

Gray indicates the differences are not statistically significant by gender at the 0.002 level.

Figure 30: Percentages of respondents indicating their work or career did not changes significantly because they became a parent by discipline.

A Closer Look at Gender Differences in Discrimination and Career Choices and Progress as a Parent by Level of Development

Finally, we ran multivariate models for countries with higher and lower levels of economic development. We find men and women from countries with higher development levels respond differently to the questions about whether their career influenced decisions about their personal life; the same is not found in countries with lower levels of development (see Figure 32, p. 76). There is a statistically significant difference in the proportion of men and women indicating their work or career did not change significantly after becoming a parent (see Figure 33, p. 77).



Each of these statistically significant relative differences (74% for men to 47% for women, or ~ 1.6 times, in Africa, for example) understates the relative likelihood between women and men in the multivariate model which accounts for confounding factors including age, employment sector, discipline, and level of development.

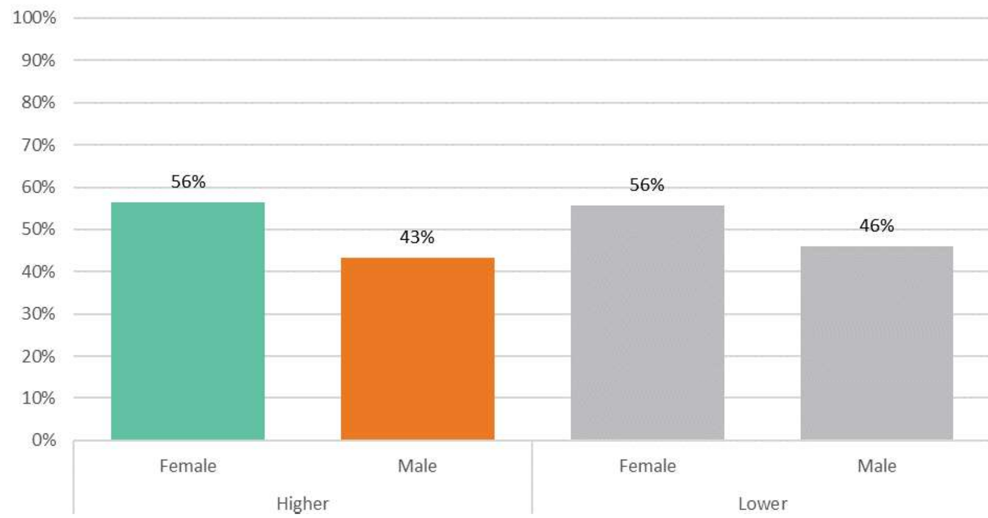
Figure 31: Percentages of respondents indicating their work or career did not change significantly after becoming a parent by geographic region.

3.6 Multivariate Analyses: Potential Areas to Examine

While our previous sections have focused on gender differences, we also used multivariate models to examine differences by discipline and by employment sector for potential lessons to be learned. We do not discuss all the differences we found below. Instead, we focus on areas where there are disciplines or employment sectors with more positive experiences than others. In those cases, there may be lessons to be learned. The analyses compare the differences for all respondents by discipline and by employment sector.

Selected Comparisons by Discipline

We tested for differences among responses by discipline after accounting for potential confounding factors including age, gender, employment sector, geographic region, and HDI. Selected results are highlighted in Table 6 (p. 78).



The statistically significant relative difference (56% for men to 43% for women, or ~ 1.3 times, in countries with higher levels of development, for example) understates the relative likelihood between women and men in the multivariate model which accounts for confounding factors including age, employment sector, discipline, and geographic region.

Gray indicates the differences are not statistically significant by gender at the 0.002 level.

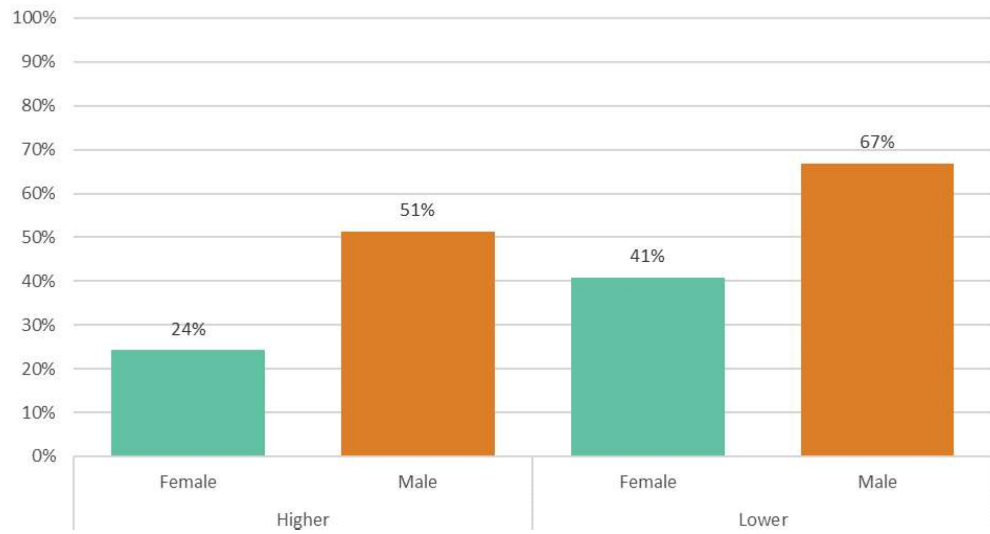
Figure 32: Percentages of respondents indicating their career influenced their decisions about children, marriage, or a similar long-term partnership.

There may be lessons to be learned about advisor support from the responses about math programs, and there may be lessons to be learned about collegial relationships from employees who studied math and physics. Finally, respondents from math and chemistry were less likely to see a change in their careers after becoming a parent.

Selected Comparisons by Employment Sector

There may be lessons to learn from different employment sectors. The different employment sectors in which respondents reported working are:

- Academics
- Government or government entity
- Industry / private sector



Both of the statistically significant relative differences (24% for women to 51% for men, or ~ 2 times, in countries with higher levels of development, for example) understate the relative likelihood between women and men in the multivariate model which accounts for confounding factors including age, employment sector, discipline, and geographic region.

Figure 33: Percentages of respondents indicating their work or career did not change significantly after becoming a parent by level of development.

- Non-governmental or non-profit organization
- Primary / secondary school
- Self-employed
- Other

We also examined differences across employment sectors. We found that respondents working in industry, non-governmental organizations and non-profits, primary / secondary schools, and respondents who were self-employed were more likely to report having respectful co-workers; those in government were less likely to do so. Respondents working in primary / secondary schools were more likely to say they had never experienced discrimination. These and other selected comparisons are highlighted in Table 7 (p. 79).

4 Discussion

These analyses provide compelling evidence that women and men do not have the same experiences in science, and that women's experiences are less positive than men's. There are gender differences in every area we examined:

- Women are more likely to report sexual harassment than men.
- Women are less likely than men to say that everyone is treated fairly in the educational system and in employment.
- Women report less positive relationships with their doctoral advisors, lower doctoral program quality, and more interruptions in doctoral studies than men.
- Women are less likely to report respectful treatment by co-workers than men. They are more likely than men to report slow career progression and discrimination.

Table 6: Disciplinary differences.

Item	Gender Differences
Doctoral advisor support†	Respondents from **Math** programs were **more likely** to rate advisor support **higher** than respondents from other disciplines.
Respectful co-workers†	Respondents whose discipline was **Physics** were **more likely** to **agree** they had respectful co-workers than respondents from other disciplines.
Never experienced discrimination†	Respondents from **Math** and **Physics** were **more likely** to indicate they had **never experienced discrimination** than respondents from other disciplines.
Promotion slowed after children†	Respondents from **Math** were **less likely** to indicate their **promotion slowed** after becoming a parent than respondents from other disciplines.
No career change after children†	Respondents from **Chemistry and Math** were **more likely** to indicate their work or career **did not change** when they became a parent than respondents form other disciplines.

These gender differences are statistically significant at the $\alpha = 0.002$ level after accounting for potential confounding factors (age, discipline, geographic region, and level of development).

† The dependent variable is ordinal, so we are unable to report a single odds ratio.

- Women have less access to career-advancing resources and opportunities than men.
- While women report better support from their families, the effect of children on women's careers is notable and not positive.
- Women are more likely than men to say that they relied on their own determination for their success in science.

Table 7: Employment sector differences.

Item	Gender Differences
Employer fair treatment†	Respondents working in **Government** were **less likely** to **agree**. Those working in **Industry, non-governmental or non-profit organizations, Primary / secondary schools** and those who were **self-employed** were **more likely** to **agree**. Respondents working in Academics and the "Other" sector were in between the two.
Respectful co-workers†	Respondents working in **Government** were **less likely** to **agree**. Those working in **Industry, non-governmental or non-profit organizations, Primary / secondary schools** and those who were **self-employed** were **more likely** to **agree**. Respondents working in Academics and the "Other" sector were in between the two.
Never experienced discrimination†	Respondents working in **Primary / secondary schools** were **more likely** to **agree** than respondents working in all other sectors.
Career influenced relationship decisions†	Respondents working in **Industry and Primary / secondary schools** were **more likely** to **agree** than respondents working in all other sectors.
Promotion slowed after children†	Respondents working in **Primary / secondary schools** were **less likely** to say their rate or promotion **slowed after children** than respondents working in all other sectors.
No career change after children†	Respondents working in **Government** and **Industry** were **more likely** to say their **career had not changed after children** than respondents working in all other sectors.

These gender differences are statistically significant at the $\alpha = 0.002$ level after accounting for potential confounding factors (age, discipline, geographic region, and level of development).

†The dependent variable is ordinal, so we are unable to report a single odds ratio.

There are other lessons to be learned. In examining doctoral programs, respondents studying in mathematics had a better perception of their advisor relationship than respondents in all other disciplines. Respondents studying computer science, mathematics, and physics were more likely to perceive they had been treated fairly in their graduate programs. Researchers can examine these programs for insight into why respondents felt this way.

In examining employment sectors, respondents working in industry, NGOs, primary/secondary schools, and who were self-employed were more likely to report having respectful co-workers than respondents working in academia. Respondents working in the government sector were least likely to report having respectful co-workers. Respondents working in industry and NGOs were more likely to report being treated fairly by their employers; this was also true for self-employed respondents. Respondents working in the government sector were less likely to report being treated fairly by their employer. Together these responses suggest that industry might have lessons to offer to other sectors on treatment of employees by co-workers and by managers.

5 Acknowledgements

The authors wish to thank Laura Merner and John Tyler for their invaluable contributions to the Global Survey of Scientists. We also wish to thank Marie-Françoise Roy and Mei-Hung Chiu for their leadership of this project. Finally, we thank all participants at the regional workshops who helped create the questionnaire.

References

[1] F. M. Dekking, C. Kraaikamp, H. P. Lopuhaä, and L. E. Meester. *A Modern Introduction to Probability and Statistics. Understanding Why and How*. London, UK: Springer, 2005. ISBN: 978-1-85233-896-1; 978-1-84996-952-9. DOI: 10.1007/1-84628-168-7.

[2] D. W. Jr. Hosmer, S. Lemeshow, and R. X. Sturdivant. *Applied Logistic Regression*. 3rd ed. New York, NY, USA: John Wiley & Sons, Inc., 2013. ISBN: 978-0-470-58247-3. DOI: 10.1002/9781118548387.

[3] R. Ivie and C. L. Tesfaye. "Women in physics: A tale of limits". In: *Physics Today* 65.2 (Feb. 2012), pp. 47–50. DOI: 10.1063/PT.3.1439. URL: https://www.aip.org/statistics/reports/global-survey-physicists.

[4] R. Ivie and S. White. "Is There a Land of Equality for Physicists? Results from the Global Survey of Physicists". In: *La Physique au Canada* 71.2 (2015), pp. 69–73.

[5] R. G. Jr. Miller. *Simultaneous Statistical Inference*. 2nd ed. New York, NY, USA: Springer, 1981. ISBN: 978-1-4613-8124-2. DOI: 10.1007/978-1-4613-8122-8.

[6] UNESCO. *The SAGA Survey of Gender Equality STI Policies and Instruments, SAGA Working Paper 3*. Paris, 2018. ISBN: 978-92-3-1002922-2. URL: https://unesdoc.unesco.org/ark:/48223/pf0000266145.