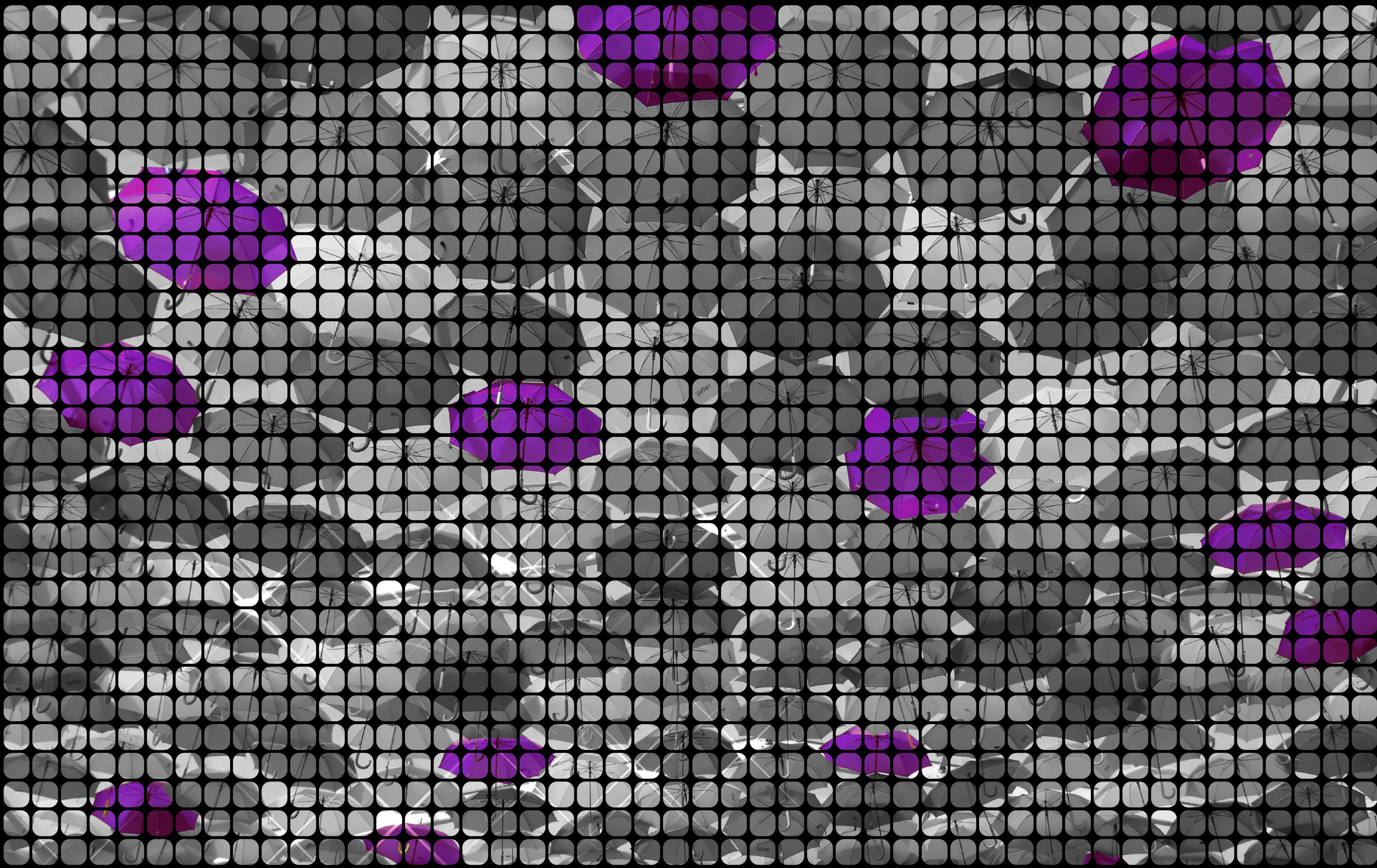

Digital Research Reports

Gender Representation in UK Research Institutions

An analysis of gender representation by field of research at UK institutions

Dr Helene Draux, Simon Porter, Ricarda Beck and Dr Suze Kundu
Foreword by James Greenwood-Lush and discussion by Dr Jessica Hamer

JANUARY 2019

DIGITAL science

Dimensions

About Digital Science

Digital Science is a technology company working to make research more efficient. We invest in, nurture and support innovative businesses and technologies that make all parts of the research process more open and effective. Our portfolio includes admired brands including Altmetric, Anywhere Access, Dimensions, Figshare, ReadCube, Symplectic, IFI Claims, GRID, Overleaf, Labguru, BioRAFT, ÜberResearch, TetraScience and Transcriptic. We believe that together, we can help researchers make a difference. Visit www.digital-science.com

About Consultancy

Our consultancy team delivers custom reporting and analysis to help you make better decisions faster. With in-depth knowledge of the historical and current research ecosystem, our unique perspective helps get the most value from data on the research lifecycle. Our team of data scientists are experts in using innovative analytical techniques to develop revealing visualisations and powerful insights. We understand the changing research landscape, and we can help you develop an evidence base on which to build the best research management and policy decisions. Visit www.digital-science.com/consultancy

About Dimensions

Dimensions is an innovative research knowledge system that re-imagines discovery and access to research. Developed by Digital Science in collaboration with over 100 leading research organizations around the world, Dimensions links grants, publications, citations, alternative metrics, clinical trials and patents. It enables users to find and access the most relevant information faster, analyze the academic and broader outcomes of research, and gather insights to inform future strategy. Data and expertise that span the research life cycle were contributed by the teams at Digital Science portfolio companies ReadCube, Altmetric, Figshare, Symplectic, Digital Science Consultancy and ÜberResearch, who came together to realize their unique strengths and share their passion for building tools that benefit the research community. Find out more at www.dimensions.ai

About the authors

Dr Hélène Draux is Research Data Scientist at Digital Science. She has an academic background in social geography, with experience in data visualisation and data science. Her work has focused on facilitating the inclusion of the general public through interactive visualisations, to communicate and collect data. She has worked in the UK and Denmark, and participated in several national and European-funded research projects. She has also worked in R&D in the private sector, where she beta-tested the use of phone applications for place discovery.
http://orcid.org/0000-0001-8837-168X

Simon Porter is Director of Innovation at Digital Science. He spent 15 years at the University of Melbourne in roles spanning the Library, Research Administration, and Information Technology. With his understanding of how information on research is collected, Simon has forged a career transforming university practices in how data about research is used, both from administrative and eResearch perspectives. In addition to making key contributions to research information, visualization and discovery within the university, Simon is well known for his advocacy of research profiling systems and their ability to create new opportunities for researchers. Simon established and ran the annual Australasian conference on research profiling. In 2012, Simon was the program chair of the annual VIVO conference.
https://orcid.org/0000-0002-6151-8423

Ricarda Beck is a Research Assistant at Digital Science Consultancy. Ricarda graduated from the University of Ulm in August 2018 with a degree in Chemistry and Business. Having moved to London from Germany in September 2018, Ricarda has been involved in research projects about the University of Cambridge, New Zealand, Austria and many others.

Dr Suze Kundu is a nanochemist by trade, and Head of Community Engagement at Digital Science. Her research area is functional semiconductor nanomaterials for applications in artificial photosynthesis. A passionate educator and collector of degrees, she has also studied for a PGCE in Senior School Science and an MEd in University Learning and Teaching. After six years lecturing at Imperial College London and the University of Surrey, she joined Digital Science in 2018 to head up engagement with all communities, whether they are academics, policy makers, school children, and anyone in between. As an experienced science communicator, she shares her love for science through live lectures, on TV and radio, and in print.

This report has been published by Digital Science which is part of Holtzbrinck, a global media company dedicated to science and education.

Digital Science, 4 Crinan Street, London N1 9XW
consultancy@digital-science.com

Foreword

By James Greenwood-Lush, Head of Athena SWAN at Advance HE
James writes in a personal capacity. Twitter: @jlush2

In order to address gender under-representation, we need to identify and locate it, undertake research and analysis to consider the underlying causes, and enact evidence-based, targeted interventions. At the heart of these processes are data. Data are the key to understanding how the design and operation of the research and higher education ecosystem manifests in terms of inequalities. They are the basis for convincing sceptical colleagues of the case for action. And they illuminate the systems and practices that – often unintentionally – perpetuate the status quo.

Data are the basis for convincing sceptical colleagues of the case for action

We need more diversity in our research system. We need it to bring different perspectives. We need it to enable everyone to contribute. We need it for social justice. We also need to understand that for things to change, deliberate endeavour is required. If we desire a more diverse system, we need to look at where we are currently failing, and face these challenges head-on.

Research organisations across the UK are making significant efforts to tackle gender inequality (many through participating in the Athena SWAN Charter) and attention and concern have become the mainstream. Institutions are recognising that identifying their challenges, changing how they operate, recognising and nurturing talent and working towards a healthier workplace culture brings benefits for diversity and the organisation as a whole. But in order to measure performance, we need to interrogate the data. People working in time-poor, pressure-rich circumstances need confidence that their efforts are making a difference. Data then give us something to build on and use to inform further improvements to our approaches and practice.

If we democratise data, we have more tools to hold other actors to account. The Digital Science interactive tool is an interesting development to support this effort and I look forward to seeing it progress. For in depth guidance, resources and information on using equality data, you can visit the Advance HE (formerly ECU) website:

https://www.ecu.ac.uk/guidance-resources/using-data-and-evidence/ and https://www.ecu.ac.uk/guidance-resources/research/

For things to change, deliberate endeavour is required

Summary

A more diverse range of researchers are better able to understand and meet the research needs of a widely varied society

This report looks at gender representation across all fields of research in UK institutions.

Gender imbalance in STEM and research in general is often discussed, and many initiatives have been put in place to attempt to redress the balance however, without a quick and easy way to analyse the scale of imbalance within a field of research and institution, and without the means to easily compare progress with others, it is difficult to identify things that are working well to nudge a change in the demographic, and to learn from and help others.

In order to make this analysis and comparison as easy as possible, we have created an interactive data visualiser that currently displays the gender splits for fields of research and UK institutions as determined by a gender guesser program that scanned all relevant *Dimensions*[1] data between 2012 and 2017, with a commitment to extending this to include other countries and updating this with new data so that changes can be monitored over time.

Using *Dimensions* data and analysis from the interactive data visualisation tool, we confirmed the gender imbalance in UK STEM research is indeed much greater than that in the Arts and Humanities, and despite some Arts and Humanities fields of research in some institutions having more women than men, a snapshot of gender splits across the whole of the UK showed that women are yet to reach gender parity in these subjects. As our interactive visualiser is updated, we hope that the richness of data and the ease of use will make it a useful tool for those monitoring gender balance in all fields of research.

Introduction

We've created an interactive data visualiser that displays the gender splits for fields of research and UK institutions

Recently, we released a report 'Gender Imbalance in Cancer Research Grants'[2] which highlighted the benefits of using the *Dimensions* database to confirm the findings of a report carried out by the BMJ showing that women were awarded fewer and smaller cancer research grants, and that the areas of research that were funded for men and women were on the whole vastly different.

We are interrogating diversity data further with this latest report as representation of all of society is important in research. A more diverse range of researchers are better able to understand and meet the research needs of a widely varied society.

It is widely reported that there is a gender imbalance in science, technology, engineering and maths (STEM), but how wide is the gender gap, and how does this compare to the arts and humanities? Have any of the initiatives implemented to address this gender imbalance been successful? In order to answer these questions, Digital Science data scientists led by Simon Porter, Director of Innovation, have created an interactive tool[3] (figure 1) to visualise the vast array of data available via the *Dimensions* database.

[1] *Dimensions* https://app.dimensions.ai/discover/publication

[2] *Gender Imbalance in Cancer Research Grants* https://doi.org/10.6084/m9.figshare.7378001.v1

[3] *Gender Representation Visualiser* https://www.digital-science.com/gender-representation-in-research-tool

Methodology

We queried *Dimensions* for publications written between 2012 to 2017, with at least one author affiliated to the UK in order to identify authors who were still working in the past 5 years. We then retrieved all authors of these publications affiliated to the UK and finally retrieved their publication history.

Table 1: Publication dataset used for the UK	
Number of publications retrieved	960,000
Number of authors with a calculated current institution in UK	302,000

We identified a little more than **300,000** researchers who had published between 2012 and 2017 and had been affiliated with a UK institution.

- **47.5%** researchers had first names that were very likely male names
- **32.6%** researchers had first names that were very likely female names
- The remaining **19.9%** could not be identified either way

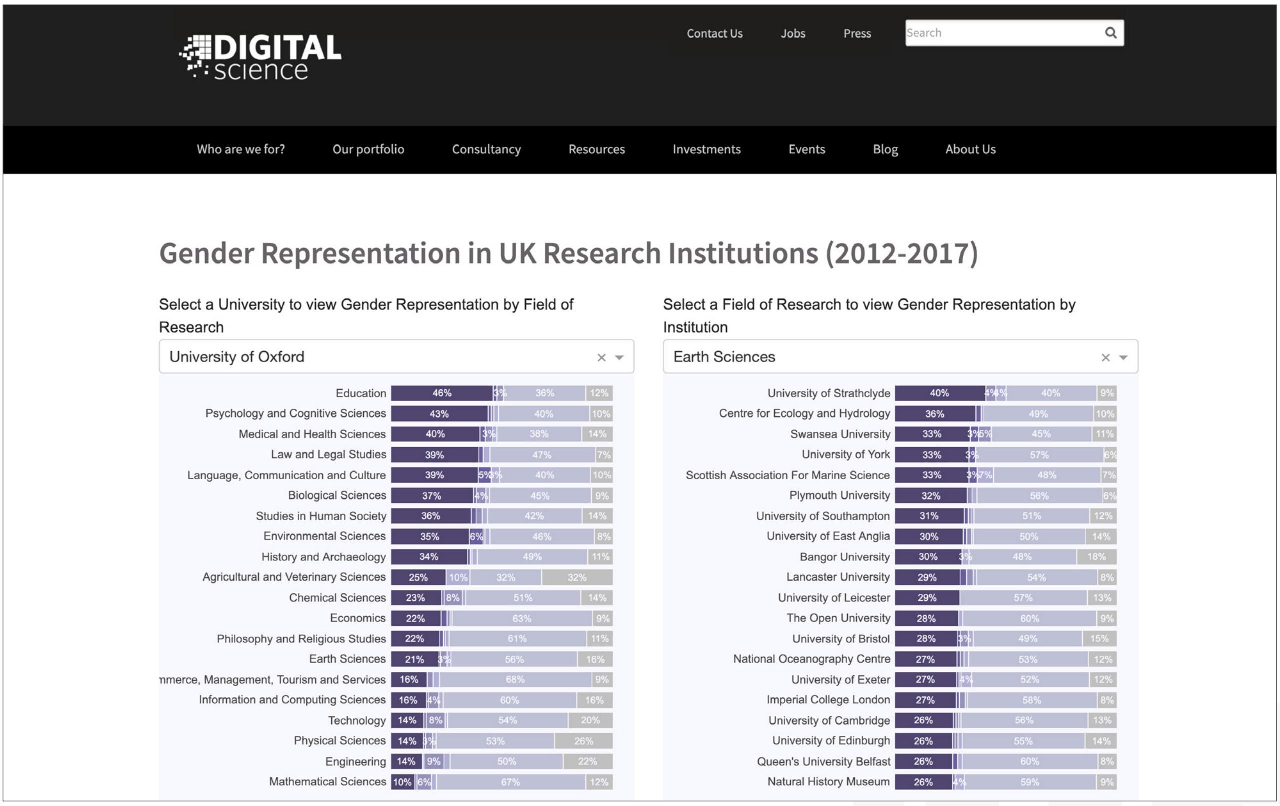

Figure 1: An interactive tool to search for gender splits across a range of fields of research for UK research institutions.

Interactive Data Visualiser

The tool shows the percentage of women researchers across all research areas of any UK institution from 2012 to 2017

Using a similar Python package gender-guesser as the one used in our first gender report, and taking into consideration all of the same limitations, the tool shows the percentage of women researchers across all research areas of any UK institution from 2012 to 2017, with plans to expand this to display institutions in other countries over the coming months, and to update this next year to see whether these figures have changed at all. The tool allows the user to search by institution or by field of research (FOR). *Dimensions* automatically classifies research published using these FOR codes, so for this study we have used the first and broadest level of research area classification to compare the proportion of each gender in these fields; for example, a materials science and engineering researcher working in a chemical engineering department would be classified under the FOR code for 'Engineering', rather than by the type of engineering or the specific strand of research. The tool is displaying gender guesser by the software as 'female name', 'mostly female name', 'androgynous name', 'mostly male name' and 'male name', with a sixth category for 'unknown'; names that the software was unable to assign a gender to.

Using the Data Visualiser

Gender splits can be searched for by institution and by field of research within each institution

Gender splits can be searched for by institution and by field of research within each institution; only FOR categories with a detected population of more than 20 researchers are shown. On the left-hand side of the website, the user can search by institution. For example, to take a look at UCL as an example of a multi-faculty university in the UK, we can either search for it in the drop-down menu on the left, or type in 'University College London', or simply 'London' and select it from the shortlist of institutions that appears in the menu (figure 2). After clicking on 'University College London' in the menu, the gender splits in the fields of research undertaken there are displayed below the menu.

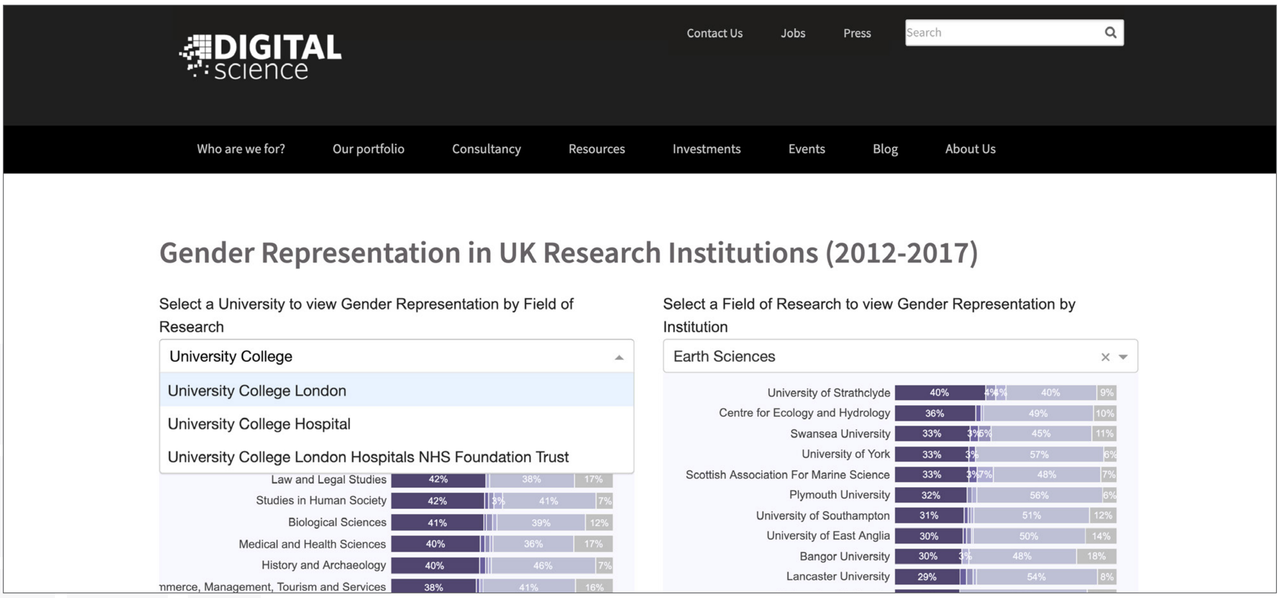

Figure 2: Using the interactive tool to search for gender splits at a specific UK research institution, e.g. UCL.

Results

We can see in figure 3a that Psychology and Cognitive Science at UCL has the highest percentage of women researchers at 49%, 34% having a male name, and 9% of unknown gender according to the gender guesser.

Following Psychology and Cognitive Science, UCL's top five fields of research for gender balance are Language and Communication also with 49%, Education with 47% women researchers, Environmental Sciences with 44% women researchers, and Law and Legal Studies and Studies in Human Society, both with 42% women researchers.

In Education, 47% of researchers at UCL are likely to be women based on their name, and 34% of researchers are likely to be men. 9% of researchers' gender could not be guessed using the tool, and the remainder were ambiguous

Focus on Education Research at UCL

The scatter graph directly below this shown in figure 3b shows how UCL's subjects compare to the same fields of research in other institutions. Points to the far right of the scatter graph have the highest percentage of women researchers. The position of each of these data points with respect to the y axis, population percentile, is determined by the actual number of women researchers as a proportion of the total research population.

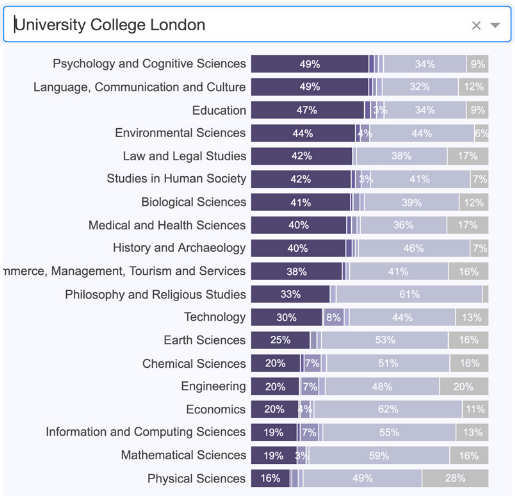

Figure 3a: Gender splits across all fields of research carried out at UCL split into gender using Dimensions data and the interactive visualiser.

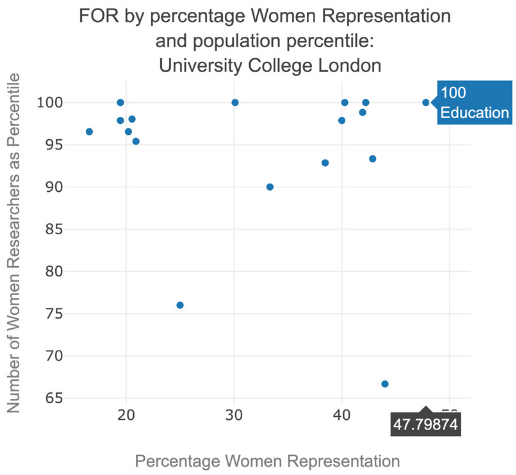

Figure 3b: Scatter graph showing the percentage of women in research across all fields of research at UCL against actual numbers of women researchers as a percentile of population.

In Education, 47% of researchers at UCL are likely to be women based on their name, and 34% of researchers are likely to be men. 9% of researchers' gender could not be guessed using the tool, and the remainder were ambiguous. Focusing on the data point for Education at UCL, we see that it is in the 100th percentile of Education research institutions in the UK, meaning that no institutions carrying out research in the field of Education have a greater number of women than UCL, and therefore all institutions have fewer women than UCL. Moving left, we see that UCL has the largest number of women of all institutions for another four fields of research, each holding the top spot at the 100th percentile, with Biological Sciences trailing ever so slightly in the 98.9th percentile.

If we continue to focus on Education, we can see which institutions have a better representation of women. This can be investigated in one of two ways; first, if we click on the Education data point in the scatter graph on the left, the data visualised on the right hand side updates to show the gender splits of institutions carrying out research in this area as shown in figure 3c. Alternatively, we can search for Education by typing this into the drop down menu. In total there are 65 UK institutions carrying out research in Education as classified in the *Dimensions* dataset. Of those 65, 27 institutions have a higher percentage of women researchers than UCL in the reporting time period, even though in numbers UCL has the most, denoted by its height against the y axis showing population percentile. 37 institutions have a lower percentage of women researchers.

From this we can see that UCL has the highest number of women researchers in Education as identified by the gender guesser (figure 3d), with the Open University in second place and the University of Cambridge in third place, however each of these institutions have higher percentage representation of women, with UCL at 48% women researchers, the University of Cambridge on 52%, and the Open University on 54%. Institutions with the highest percentage representation of women in Education research are the Glasgow Caledonian University with 62% women researchers, followed by Robert Gordon University, Brunel University London, the University of York and Bournemouth University all with 62% as shown in figure 3b.

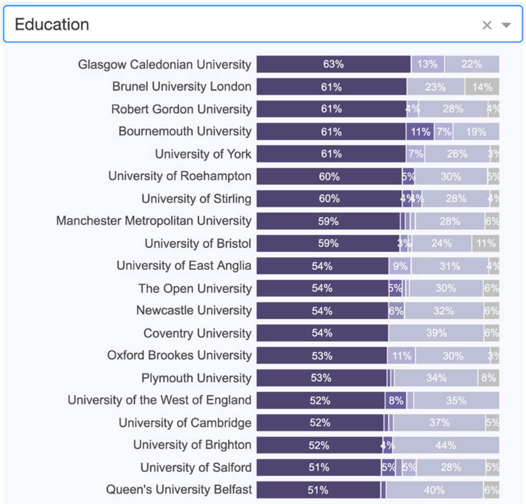

Figure 3c: Chart showing the percentage of women in Education research across all UK institutions.

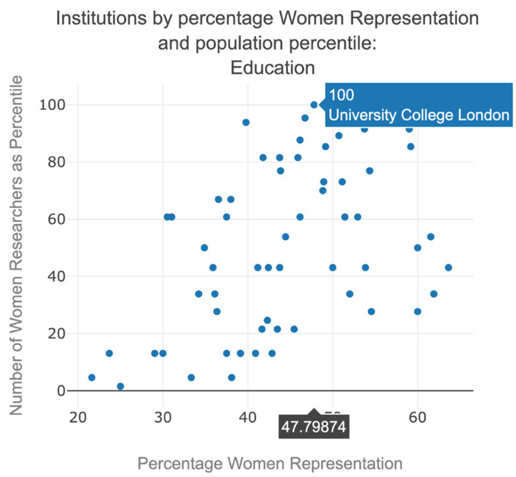

Figure 3d: Scatter graph showing the percentage of women in Education research across all UK institutions against actual numbers of women researchers as a percentile of population.

Overall it can be said that Education has a good representation of women researchers (figure 3c), with women outnumbering men in more than twenty research institutions within this field of research. The distribution of institutions shown in figure 3e does however indicate that not all Education research institutions have more women than men, as the histogram tails off to the left with 20 institutions below 40% representation.

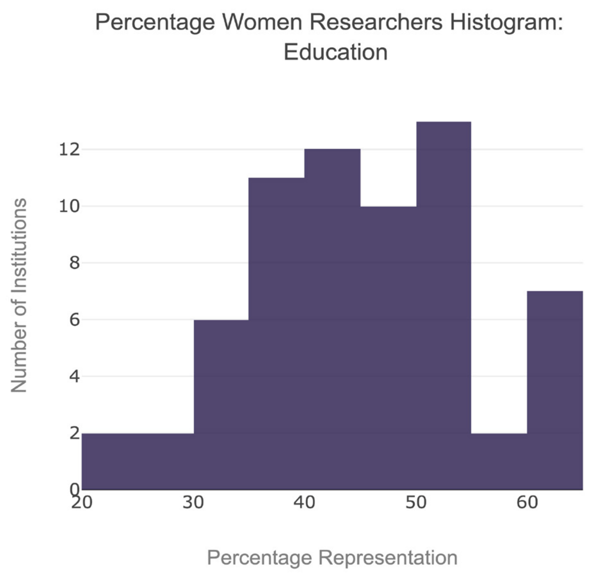

Figure 3e: Histogram showing distribution of percentage of women researchers in Education across all UK institutions.

STEM Research vs Arts and Humanities Research at UCL

On the other end of the scale, and to once again use UCL as an example, the subjects with the lowest percentage of women researchers were Physical Sciences with just 17% women researchers, and Mathematical Sciences and Information and Computing Systems each with 19% women researchers. In fact, generally speaking, there seems to be a split between Science, Technology, Engineering and Maths (STEM) subjects which have a lower representation of women researchers, and the Arts and Humanities which have a higher representation of women researchers, though many of these subjects still tend to veer far from gender parity, and go against the assumption that Arts and Humanities research is dominated by women.

Many of these subjects still tend to veer far from gender parity, and go against the assumption that Arts and Humanities research is dominated by women

Gender Balance in UK Research

UCL's gender imbalance is largely in line with UK trends. Across the UK, representation of researchers with strongly type female names is only greater than male names in Psychology and Cognitive Sciences, Language, Communication and Culture, and Education (figure 4).

It is clear that women are under-represented in the STEM subjects when compared to Arts and Humanities. Biological Sciences and Medical and Health Sciences slightly buck this trend, coming 6th and 7th overall in the 22 fields of research assigned.

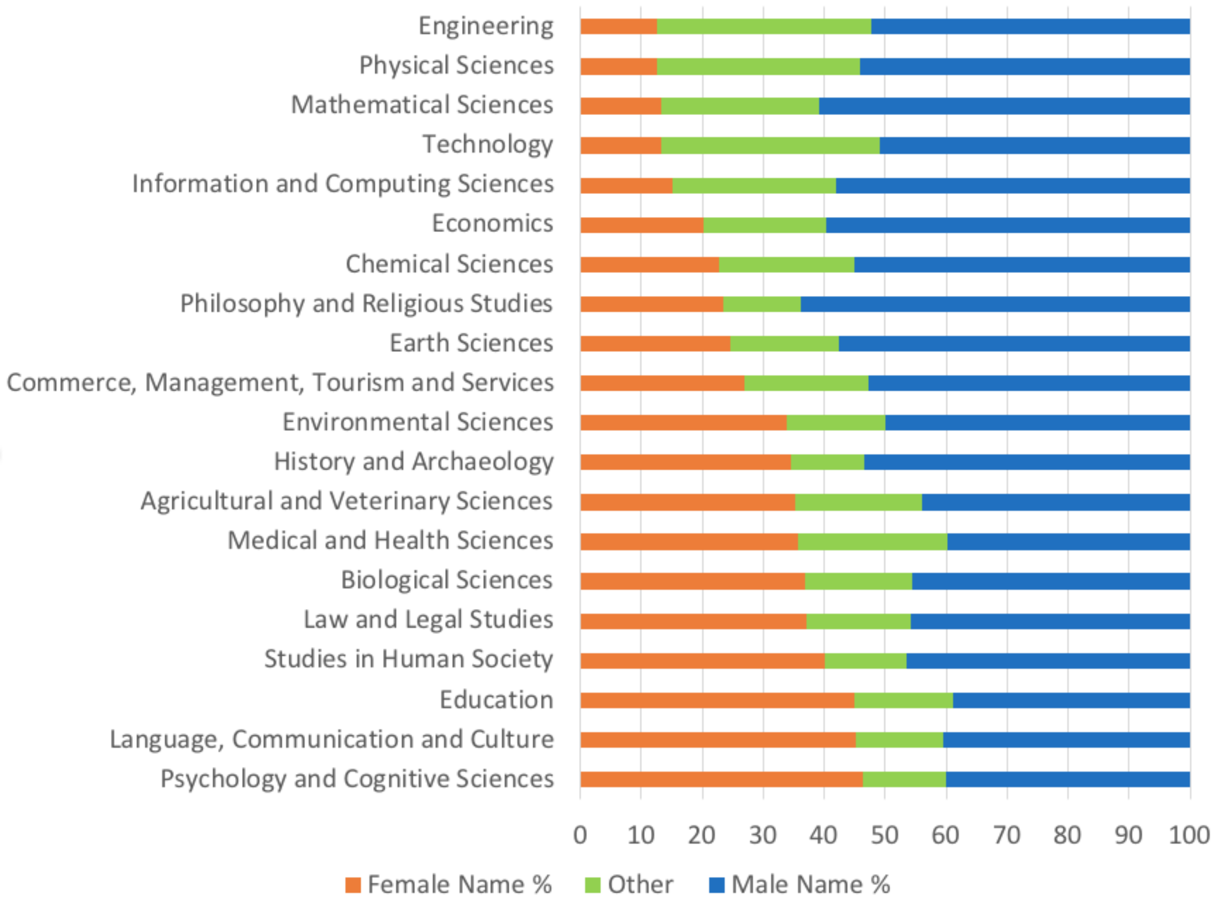

Figure 4: Gender split across all fields of research at UK research institutions arranged by percentage of women researchers

There is also an interesting trend in the percentage of researchers that could not be identified by the gender guesser, as this seems to increase as the percentage of women decreases, and as we move from the Arts and Humanities towards the STEM subjects. This could be related to the broader ethnic diversity of researchers working in these fields, and is something that we would like to investigate in our following report, in order to improve our visualisation tool and make it as useful and inclusive as possible, reducing any potential western bias in the results.

STEM Research vs Arts and Humanities Research Across the UK

Looking at the distribution of percentages of women in different fields of research (figures 5a – 5f), STEM subjects such as Physical Sciences, Technology and Information and Computing Sciences display distributions across all institutions that do not extend beyond more than 40% women, with both Technology and Physical Sciences showing peak representation of women in this field of research as between 5% and 15%. On the other hand, fields such as Education and Language, Communication and Culture (not pictured) have peak representation around 50% or higher, with a narrow distribution between around 30% to 60%. Some subjects such as Medical and Health Science and Psychology and Cognitive Sciences have a huge distribution range of representation of women within their fields of research, with the former spanning zero representation of women all the way up to 70% representation.

Both Technology and Physical Sciences show peak representation of women in this field of research as between 5% and 15%

Education and Language, Communication and Culture have peak representation around 50% or higher

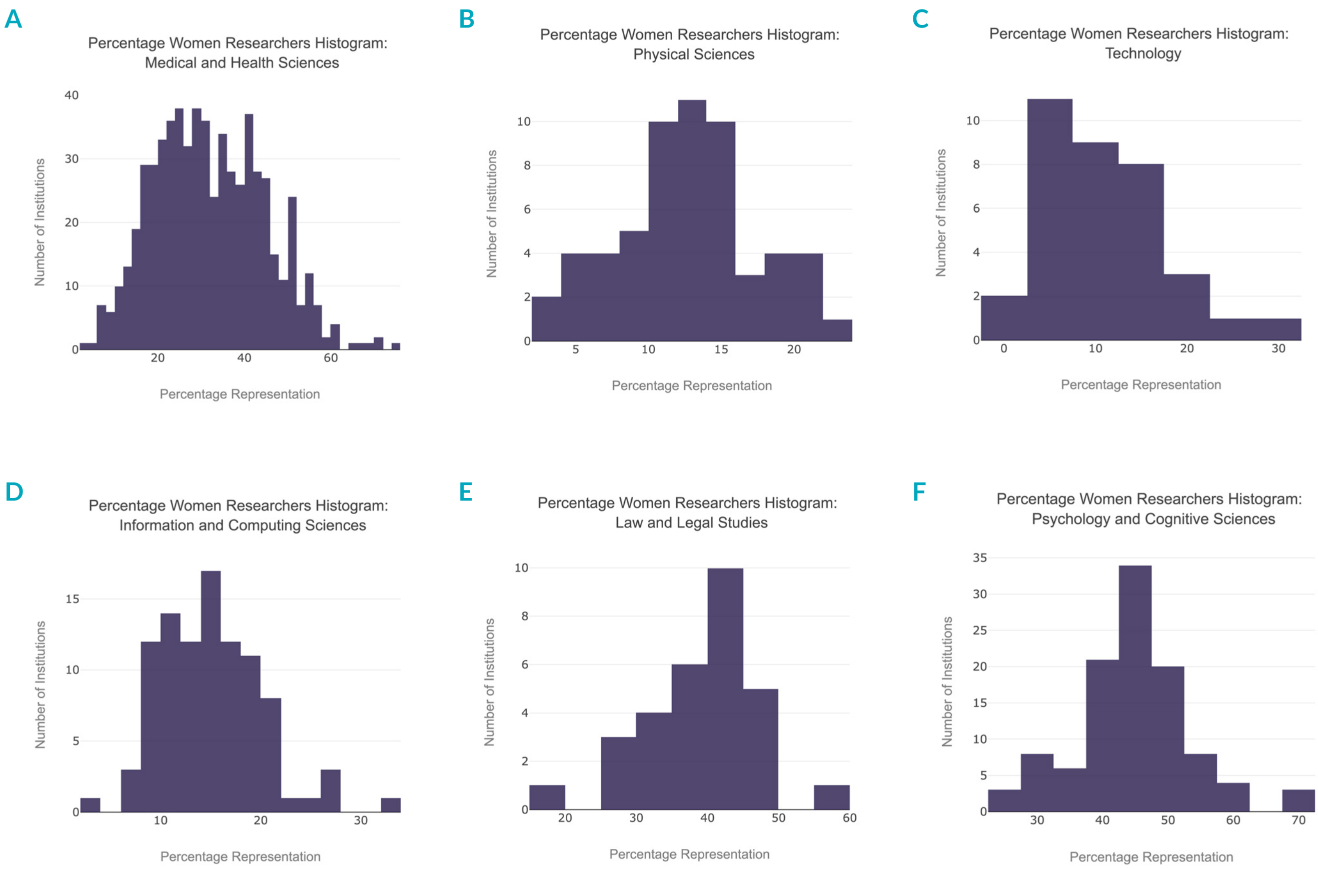

Figure 5: Distribution of percentage representation across six examples of fields of research across STEM and the Arts and Humanities (from top left to bottom right: Medical and Health Sciences, Physical Sciences, Technology, Information and Computing Sciences, Law and Legal Studies, and Psychology and Cognitive Sciences).

Conclusion

These figures confirm that there is a problem with gender imbalance in STEM subjects and also show that Arts and Humanities subjects do not have a majority of women researchers as is often presumed

By analysing the wealth of data within the *Dimensions* database with this Python package that can seek out authors of publications between set time periods and guess their gender by their name, we have created a comprehensive snapshot of the gender demographics of the research landscape which can be viewed nationally, or on an institutional or field of research level. The differences within each field of research confirmed that women are less represented in STEM than in the Arts and Humanities, but also that these fields of research have not yet reached gender parity themselves.

The current iteration of the interactive tool provides a quick and easy way to visualise a large amount of data. These figures are based on researchers active between 2012 and 2017, and confirm that there is a problem with gender imbalance in STEM subjects, however it also shows that Arts and Humanities subjects do not have a majority of women researchers as is often presumed.

Over the coming months we will be releasing interactive tools for other countries, and will be updating the data each year in order to provide researchers, funders and institutions with a quick and easy way to monitor changes in their research demographic, and find out which initiatives are successfully working to create a more diverse and inclusive set of minds to tackle the wide range of challenges researchers are faced with. We will also be improving the tool to ensure that it is as inclusive as possible, with an aim to reduce the number of researchers whose gender could not be determined by the gender guesser tool, which could be indicative of a slight western bias in the results that the tool can currently display.

These studies show the importance of having disambiguated researchers and well categorised data in order to conduct meaningful studies. Using *Dimensions*, we revealed some interesting insights into the true gender imbalance across institutions and across a range of subjects, and showcased the use of an interactive tool created to help easily visualise this data, with a commitment to updating this data in the future to aid monitoring of the success of various schemes and interventions.

Discussion

By Jessica Hamer, Education Consultant and Physics Coach
Twitter @DrJessicaHamer

In the UK, there is a gender imbalance in STEM (science, technology, engineering and maths) from the classroom through to academia and industry – that is fact. Despite increases in the number of women choosing to study subjects such as medicine and biology over the last 50 years, women are still underrepresented in most of the STEM subjects, and academia and the higher echelons of industry more broadly. Too many students and professionals are still turning their backs on rewarding learning and careers simply because they feel they don't belong; that there are invisible barriers within the social construct of their institution and wider society that prevents them from remaining and/or progressing in STEM. And it's the workplace that suffers as a consequence; with increased diversity, teams of STEM professionals would have increased innovation and discovery and ultimately greater financial results[1,2].

Women and girls can "opt-out" of STEM at a young age. When it comes to exams, boys are more than three times less likely to sit an A-level psychology exam and yet are more than three times more likely to choose physics A-level than girls. Is this surprising when research from the Institute of Engineering and Technology (IET) found that girls were three times less likely to be given a STEM toy for Christmas? Correlation does not equal causation, but when co-educational schools nationally have proportionally fewer girls continuing physics and boys continuing in English beyond GCSE compared to single-sex schools you have to ask: what is going on within the school environment to make this happen?[3]. There is limited research on this topic, but what is out there suggests that it is complex issue formed out of gender stereotypes, school systems, unconscious (and in some instances outright conscious) bias in the classroom and student perceptions of the STEM subjects and the opportunities it could provide them with in the future.

We have to ask: what is going on within the school environment to make this happen?

Promisingly, the total number of women selecting a degree in STEM has increased year on year, indicating that there is a reserve of women resilient enough to overcome these invisible barriers. However, according to WISE (Women in Science and Engineering), women still only make up 23% of the STEM workforce, if looking at engineering in isolation that number drops to 12%. Retention and progression of women in many organisations and institutions appears to be an issue. A report published in 2018 by the Royal Society found that for those awarded a Dorothy Hodgkin Fellowship or a University Research Fellowship, it took women on average more than a year longer to achieve a chair or other senior position. Indeed, in chemistry women hold just 9% of UK chemistry professorships meaning that after

[1] *Nielsen, M.W., Alegria, S., Börjeson, L., Etzkowitz, H., Falk-Krzesinski, H., Joshi, A., Leahey, E., Smith-Doerr, L., Williams Woolley, A., and Schiebinger L., (2017), PNAS 114 (8) 1740-1742*

[2] *McKinsey and Company (2018) Delivering through Diversity*

[3] *Institute of Physics (2013) Closing Doors: Exploring gender and subject choice in schools*

Research, such as that contained within this report, helps us to identify and break the barriers that prevent women from progressing, and are therefore invaluable in ensuring lasting change

undergraduate level, the relative proportion of female chemists drops by 35%[4]. The reasons why women choose to leave STEM or decide not to move forward in their career are numerous and are likely to vary widely between academia and industry, but may include pressures from short-term contracts, long hours, a lack of transparency in relation to promotion, lack of flexible working options and workplace culture.

With the rise of the #MeToo movement there has been a gear change in the national mood in relation to gender equality, but in STEM any real change feels glacially slow. The causes of the gender imbalance in STEM are myriad and as a society we are only beginning to look into understanding and addressing these. Research, such as that contained within this report, helps us to identify and break the barriers that prevent women from progressing, and are therefore invaluable in ensuring lasting change.

[4] *Royal Society of Chemistry (2018) Breaking the Barriers*

Part of **DIGITAL**science

Altmetric	ANYWHERE ACCESS
BIORAFT	Dimensions
figshare	GRID
ifi CLAIMS	labguru
Overleaf	readcube
SYMPLECTIC	TETRASCIENCE
transcriptic	über RESEARCH

DIGITALscience Consultancy

digital-science.com