Data, demographics, and diversity: improving the quality of evidence and reporting on representation in the STEM workforce

As part of the inquiry into equity in the science, technology, engineering, and maths (STEM) workforce, on 18 February 2021, the All-Party Parliamentary Group for Diversity and Inclusion in STEM held the second of four evidence roundtables. Titled *Data, demographics, and diversity: improving the quality of evidence and reporting on representation in the STEM workforce,* this closed roundtable convened experts in equality, diversity & inclusion (EDI) and data from across STEM sectors to discuss underrepresented communities within STEM.

The roundtable sought to collect expert insight on how to combat inequity within the STEM workforce and facilitate sharing of professional knowledge to improve the quality of evidence and data reporting on representation in the STEM workforce.

The views expressed and potential recommendations discussed have been recorded as contributing evidence to the inquiry and will help shape the final report, due for publication in summer 2021.

The roundtable used the following questions to prompt the discussion:

1. How can we improve analysis and reporting of Protected Characteristics (or similar) in an intersectional respectful manner?
   - How can we combat issues around self-identification?
2. What STEM data are we missing, that could be reported on, and what sets could be utilised?
   - Can we break this data down by subsector?
3. What changes can be made to national data sets to improve reporting on the STEM workforce?
4. How important do you think better reporting/more detailed data on Protected Characteristics is to your organisation(s)/workforce?

The following points were highlighted, many of which could be applied to the whole workforce, not just STEM.

Data collection

Employers, organisations and the Government need to collect data not only on the issues surrounding representation and retention in STEM but also on what interventions have been tried to combat these issues, along with meaningful assessments of these interventions. We need better theoretical models of how positive changes will be evidenced in data employers and organisations can collect. UKRI has published two reviews on the effectiveness of EDI interventions, for the UK and internationally.

Qualitative data is needed to investigate the narratives of how exclusion occurs and effectively overcome barriers. We shouldn't rely on quantitative data alone; qualitative data is essential to help show and understand experiences different individuals and communities face in the STEM workforce.

The Inclusive Data Task Force has been launched by the National Statistician and is looking at how data is collected, analysed and disseminated. There is a tension between harmonization of statistics on the one hand and allowing people space for self-identification on the other. A public consultation is currently open. The Census will be taking place in March and will gather greater information on gender and sexual identity than before.

There is currently no agreed definition of the STEM workforce within data collection. The Royal Society is working with the Institute of Employment Research at the University of Warwick to develop a methodology for defining the STEM workforce using the ONS Standard Occupation Codes. The Society hopes to publish the methodology in the coming months

The majority of the STEM workforce work in SMEs, not academia or large companies. They have different constraints when it comes to data collection - it needs to be achievable, practical and standardised for all employers. The Mayor of London Workforce Data Equality Guide contains information on how to collect, analyse and act on equalities. Longitudinal career information would be useful to help employers, networks and Government understand pipelines and progression.

Demographic diversity data on the workforce should be more granular and intersectional. One example is the use of the problematic term 'BAME', it groups together anyone who isn't white as 'other'. Society needs to move away from its usage as a homogenising term which obscures the experiences of Black people, Asian people and those from Ethnic Minority backgrounds, under its umbrella. Similarly, all LGBTQ+ people cannot be grouped together, representation in STEM varies between communities.

The priorities should be to collect qualitative narratives, use more sophisticated data analysis and mining methods and make use of methodologies for measuring inclusions and culture.

Self-identification/reporting and trust

Demographic data collection in the workforce is often hampered by low rates of self-identification and reporting. There is often a lack of trust in how employers or the Government will use data, which impacts willingness to respond. Many employees do not believe data collected will actually be used, others who have never experienced difficulties in the workplace don't bother as they don't think it applies to them.

Employees may not want to report things such as a disability for fear it will harm their career or may not respond out of fear of identification. Employees often have to learn to hide aspects of their identity in the workplace. There needs to be an increase in representation and visibility of underrepresented employees in the workplace, so employees and applicants can identify with others like themselves.

The language used to describe disability and the perception of it in society also impacts how people respond and self-report. For example, someone may identify with having a disability such as dyslexia or depression, but not consider themselves disabled or even be aware what constitutes a disability. The Royal Society report on disability STEM data for students and staff in higher education found that the proportion of disabled students in HE has increased, but this is driven by a significant increase in the proportion of students declaring mental health conditions and learning differences. The proportion of students with physical disabilities has not increased significantly.

The options given for self-identification of ethnicity also create problems, as mentioned above, the term 'BAME' is problematic and obscuring. Tick-box options are limiting for anyone who doesn't fall neatly into one category, or for example who may be from a mixed background but identifies as 'British'.

The Physiological Society found that younger generations were more willing to participate in data collection, which can make data sets biased.

Videos accompanying surveys or questionnaires to illustrate why it is important to respond, help to improve participation rates, these have been used by the NHS. These explain how the data will be used and increase trust levels by humanising it.

Socio-economic status and geography

Socioeconomic status is an important part of the intersectional diversity picture and one that is not easy to measure and track as it is not static. There is also a tendency of employees in certain professions to report their background as of lower socioeconomic status than is accurate.

There is also geographical disadvantage or advantage in the UK, based on where a person lives geographically. Areas such as the 'golden triangle' (Cambridge, Oxford and London) provide different opportunities to the North East, for example. Geographical origin data is not always collected but can greatly impact careers and success.

Safe spaces, reverse mentoring and inclusion

Psychological safety, as explained by Amy Edmondson at Harvard University, is the seed for inclusion, you cannot have an inclusive workforce without it. Creating safe spaces for harder conversations to educate the workforce on diversity are important. Open dialogue and sharing of experiences, as well as showcasing a variety of role models in the workforce are also important. Examples of these actions can be found in Deloitte's D+I strategy.

Reverse mentoring helps senior management to understand and appreciate the actual issues impacting staff in a workplace, and what's happening on the ground. Senior individuals need to be receptive for this to work. Reverse mentoring is included in the NHS Workforce Race Equality Standard 15 year model.

The phrase 'nothing about us without us' should be applicable to all inclusion work, those who are impacted need to be consulted. However, it is also important that this work doesn't fall solely on the shoulders of these underrepresented groups.

The roundtable was chaired by Baroness Brown of Cambridge and attended by the following participants:

- **Katy Amberley**
  Chief Executive, The British Society for Haematology
- **Nadine Dyer**
  Financial Advisory, Respect and Inclusion Lead at Deloitte
- **Lori Frecker**
  Diversity and Inclusion Programme Manager, the Royal Society
- **Veronica van Heyningen CBE FMedSci FRS**
  Diversity Committee Chair, the Royal Society
- **Dr Eugenie Hunsicker**
  Director of Equality and Diversity, School of Science; Senior Lecturer at Loughborough University
- **Payal Jain**
  Chair, Women in data
- **Liz McKeown**
  Director, Public Policy Analysis, ONS
- **Dr Beth Montague-Hellen**
  Senior Research Librarian at the University of Nottingham, Founder of LGBTSTEM
- **Dr Keith Siew**
  Sir Henry Wellcome Postdoctoral Fellow, Scientific Editor at The Physiological Society
- **Surash Surash**
  Consultant Neurosurgeon at the RVI in Newcastle
- **Professor Tom Welton OBE DPhil CChem FRSC**
  President, Royal Society of Chemistry

This roundtable was the second of four closed roundtable sessions, the closed format was chosen to allow participants to share their experiences more freely. The roundtables follow a public launch event at which a panel of speakers addressed key issues arising from the Data Analysis Brief. The evidence gathered in these sessions supplements the open Call for Evidence, which closed January 29 2021 and resulted in over 80 submissions from across the STEM sectors.